PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, 5/10, C07K 14/705, 16/28, C12Q 1/68, G01N 33/50, A61K 38/17

A2

(11) International Publication Number: WO 00/29583

(43) International Publication Date: 25 May 2000 (25.05.00)

(21) International Application Number: PCT/US99/27566

(22) International Filing Date: 19 November 1999 (19.11.99)

(30) Priority Data:

Not furnished	19 November 1998 (19.11.98)	US
60/113,635	22 December 1998 (22.12.98)	US
60/128,194	7 April 1999 (07.04.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/113,635 (CIP)
Filed on	22 December 1998 (22.12.98)
US	60/128,194 (CIP)
Filed on	7 April 1999 (07.04.99)
US	Not furnished (CIP)
Filed on	19 November 1998 (19.11.98)
US	09/195,853 (CIP)
Filed on	19 November 1998 (19.11.98)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). YANG, Junming [CN/US]; 7136 Clarendon Street, San Jose, CA 95129 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: IMMUNOGLOBULIN SUPERFAMILY PROTEINS

(57) Abstract

The invention provides human immunoglobulin superfamily proteins (IGFAM) and polynucleotides which identify and encode IGFAM. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of IGFAM.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IMMUNOGLOBULIN SUPERFAMILY PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of immunoglobulin superfamily proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, immune system disorders, and infections.

BACKGROUND OF THE INVENTION

All vertebrates have developed sophisticated and complex immune systems that provide protection from viral, bacterial, fungal, and parasitic infections. Protection is mediated through cell surface and soluble molecules which function in recognition, adhesion or binding. The vertebrate immune system evolved from a common evolutionary precursor (i.e., these proteins have structural homology). A number of molecules outside the immune system that have similar functions are also derived from this same evolutionary precursor.

An important characteristic of the immune system is its ability to recognize and destroy foreign molecules, or antigens. Antigen recognition is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as granulocytes, monocytes, and lymphocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily. The cell surface and soluble molecules of the immune system are classified as members of the Ig superfamily, members of which contain one or more repeats of a conserved structural Ig domain. The Ig domain, 70-110 amino acid residues in length, is homologous to either Ig variable-like (V) or Ig constant-like (C) domains. The Ig domain is described as antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors, neural cell adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR).

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of β-sheets. Conserved cysteine residues form an intrachain disulfide-bonded loop, 55-75 amino acid residues in length, which connects the two layers of the β-sheets. Each β-sheet has three or four anti-parallel β-strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the β-strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an

additional pair of β-strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead attached to the plasma membrane by a covalent glycophosphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody and pairing of variable regions of light and heavy chains forms the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp.142-145.)

Antibodies

Antibodies, or immunoglobulins, are the founding members of the Ig superfamily and are the central components of the humoral immune response. Antibodies are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize blood-borne foreign antigens. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ε, γ, and μ H-chain types. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class

of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pages 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain repeated Ig domains. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region. In addition, H chains such as μ have been shown to associate with other polypeptides during differentiation of the B-cell. One such polypeptide called 8HS-20 is itself a member of the Ig superfamily and contains a single Ig domain (Shirasawa, T. et al. (1993) EMBO J. 12:1827-1834).

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

Antibodies can also be described in terms of their two main functional domains. Antigen recognition is mediated by the Fab (antigen binding fragment) region of the antibody, while effector functions are mediated by the Fc (crystallizable fragment) region. Binding of antibody to an antigen, such as a bacterium, triggers the destruction of the antigen by phagocytic white blood cells such as macrophages and neutrophils. These cells express surface receptors that specifically bind to the

antibody Fc region and allow the phagocytic cells to engulf, ingest, and degrade the antibody-bound antigen. The Fc receptors expressed by phagocytic cells are single-pass transmembrane glycoproteins of about 300 to 400 amino acids (Sears, D. W. et al. (1990) J. Immunol. 144:371-378). The extracellular portion of the Fc receptor typically contains two or three Ig domains.

Unique variants of Fc receptors have been identified in myeloid and lymphoid cells (Samaridis, J. and Colonna, M. (1997) Eur. J. Immunol. 27:660-665). Like typical Fc receptors, these proteins contain extracellular Ig domains and are encoded by cDNAs designated ILT1 and ILT2 (Ig-like transcripts 1 and 2). However, the transmembrane and cytoplasmic domains diverge significantly. In particular, the cytoplasmic domain is extended and contains protein motifs consistent with a role in intracellular signal transduction.

A new member of the Ig superfamily appears to play a structural role in the control of monocyte migration across epithelium or endothelium to sites of inflammation. This protein, called junctional adhesion molecule (JAM), is situated at tight junctions which occur between adjacent epithelial or endothelial cells (Martin-Padura, I. et al. (1998) J. Cell Biol. 142:117-127). JAM is 300 amino acids in length and contains two Ig domains. A monoclonal antibody (mAb) directed against JAM inhibited transmigration of monocytes across endothelial cell layers in vitro. Furthermore, systemic administration of this mAb to mice prevented recruitment of monocytes to sites of inflammation.

Viral proteins which contain Ig domains have also been described (Senkevich, T. G. et al. (1996) Science 273:813-816). These proteins include MHC-like homologs identified in the human tumorigenic poxvirus Molluscum contagiosum. Such proteins may provide a mechanism by which the virus evades the host's immunologic surveillance system.

T-cell receptors are both structurally and functionally related to antibodies. (Reviewed in Alberts, supra, pp. 1228-1229.) T-cell receptors are cell surface proteins that bind foreign antigens and mediate diverse aspects of the immune response. A typical T-cell receptor is a heterodimer comprised of two disulfide-linked polypeptide chains called α and β. Each chain is about 280 amino acids in length and contains one variable region and one constant region. Each variable or constant region folds into an Ig domain. The variable regions from the α and β chains come together in the heterodimer to form the antigen recognition site. T-cell receptor diversity is generated by somatic rearrangement of gene segments encoding the α and β chains. T-cell receptors recognize small peptide antigens that are expressed on the surface of antigen-presenting cells and pathogen-infected cells. These peptide antigens are presented on the cell surface in association with major histocompatibility proteins which provide the proper context for antigen recognition.

Synaptic Membrane Glycoproteins

Specialized cell junctions can occur at points of cell-cell contact. Among these cell junctions are communicating junctions which mediate the passage of chemical and electrical signals between cells. In the central nervous system, communicating junctions between neurons are known as synaptic junctions. They are composed of the membranes and cytoskeletons of the pre- and post-synaptic neurons. Some glycoproteins, found in biochemically isolated synaptic subfractions such as the synaptic membrane (SM) and postsynaptic density (PSD) fractions, have been identified and their functions established. An example is the SM glycoprotein, gp50, identified as the β2 subunit of the Na^+/K^+-ATPase.

Glycoproteins in the SM and PSD which have their oligosaccharide domains facing the synaptic junction are in a good position to mediate adhesive interactions between neurons. The PAC 1 glycoproteins , components of the PSD, have been identified as members of the cadherin family, proteins involved in the Ca^{2+}-dependent cell-cell adhesion in vertebrate tissues. Further support of these molecules mediating adhesive interactions is the presence of integrin-type adhesion molecules and NCAM, a member of the Ig superfamily, in the SM.

Two glycoproteins, gp65 and gp55, are major components of synaptic membranes prepared from rat forebrain. They are members of the Ig superfamily containing three and two Ig domains, respectively. As members of the Ig superfamily, it is proposed that a possible function of these proteins is to mediate adhesive interactions at the synaptic junction. (Langnaese, K. et al. (1997) J. Biol. Chem. 272(2):821-827.)

The discovery of new immunoglobulin superfamily proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, immune system disorders, and infections.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, immunoglobulin superfamily proteins, referred to collectively as "IGFAM" and individually as "IGFAM-1," "IGFAM-2," "IGFAM-3," "IGFAM-4," "IGFAM-5," "IGFAM-6," "IGFAM-7," "IGFAM-8," "IGFAM-9," "IGFAM-10," "IGFAM-11," "IGFAM-12," "IGFAM-13," "IGFAM-14," "IGFAM-15," "IGFAM-16," "IGFAM-17," "IGFAM-18," and "IGFAM-19." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-19.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-19 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the

invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-19 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of IGFAM, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of IGFAM, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding IGFAM.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of IGFAM.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding IGFAM were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze IGFAM, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"IGFAM" refers to the amino acid sequences of substantially purified IGFAM obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of IGFAM. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of IGFAM either by directly interacting with IGFAM or by acting on components of the biological pathway in which IGFAM participates.

An "allelic variant" is an alternative form of the gene encoding IGFAM. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding IGFAM include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as IGFAM or a polypeptide with at least one functional characteristic of IGFAM. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding IGFAM, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding IGFAM. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent IGFAM. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of IGFAM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of IGFAM. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of IGFAM either by directly interacting with IGFAM or by acting on components of the biological pathway in which IGFAM participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind IGFAM polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or

oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic IGFAM, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding IGFAM or fragments of IGFAM may

be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs and/or cDNAs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of IGFAM or the polynucleotide encoding IGFAM which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:20-38 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:20-38, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:20-38 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:20-38 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:20-38 and the region of SEQ ID NO:20-38 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-19 is encoded by a fragment of SEQ ID NO:20-38. A fragment of SEQ ID NO:1-19 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-19. For example, a fragment of SEQ ID NO:1-19 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-19. The precise length of a fragment of SEQ ID NO:1-19 and the region of SEQ ID NO:1-19 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis

programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default

parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding

between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune

disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of IGFAM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of IGFAM.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding IGFAM, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,

or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence.

This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding IGFAM, or fragments thereof, or IGFAM itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic

acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human immunoglobulin superfamily proteins (IGFAM), the polynucleotides encoding IGFAM, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, immune system disorders, and infections.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding IGFAM. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each IGFAM were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each IGFAM and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs. Note that in column 5, all but one of the polypeptides of the invention contain one or more Ig domains as predicted by protein function analysis programs such as PROFILESCAN, BLIMPS, PFAM, and MOTIFS. The polypeptide of the invention which lacks a predicted Ig domain (SEQ ID NO:4) does, however, show significant similarity with the viral Ig-containing protein MC51L-53L-54L.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding IGFAM. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:20-38 and to distinguish between SEQ ID NO:20-38 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express IGFAM as a fraction of total tissues expressing IGFAM. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing IGFAM as a fraction of total tissues expressing IGFAM. Column 5 lists the vectors used to subclone each cDNA library. Note that the nucleotide sequences of SEQ ID NO:23, and SEQ ID NO:27 are expressed primarily in cells and tissues associated with the hematopoietic/immune system and inflammation.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding IGFAM were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3

shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses IGFAM variants. A preferred IGFAM variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the IGFAM amino acid sequence, and which contains at least one functional or structural characteristic of IGFAM.

The invention also encompasses polynucleotides which encode IGFAM. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:20-38, which encodes IGFAM.

The invention also encompasses a variant of a polynucleotide sequence encoding IGFAM. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding IGFAM. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:20-38 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:20-38. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of IGFAM.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding IGFAM, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring IGFAM, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode IGFAM and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring IGFAM under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding IGFAM or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding IGFAM and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a

greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode IGFAM and IGFAM derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding IGFAM or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:20-38 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding IGFAM may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids

Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode IGFAM may be cloned in recombinant DNA molecules that direct expression of IGFAM, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express IGFAM.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter IGFAM-encoding sequences for a variety of purposes including, but

not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding IGFAM may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, IGFAM itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of IGFAM, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active IGFAM, the nucleotide sequences encoding IGFAM or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding IGFAM. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding IGFAM. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding IGFAM and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used.

(See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding IGFAM and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding IGFAM. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding IGFAM. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding IGFAM can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding IGFAM into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of IGFAM are needed, e.g. for the production of antibodies, vectors which direct high level expression of IGFAM may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of IGFAM. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of IGFAM. Transcription of sequences encoding IGFAM may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding IGFAM may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses IGFAM in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of IGFAM in cell lines is preferred. For example, sequences encoding IGFAM can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et

al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding IGFAM is inserted within a marker gene sequence, transformed cells containing sequences encoding IGFAM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding IGFAM under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding IGFAM and that express IGFAM may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of IGFAM using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on IGFAM is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding IGFAM include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding IGFAM, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding IGFAM may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode IGFAM may be designed to contain signal sequences which direct secretion of IGFAM through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding IGFAM may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric IGFAM protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of IGFAM activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST),

maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the IGFAM encoding sequence and the heterologous protein sequence, so that IGFAM may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled IGFAM may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of IGFAM may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of IGFAM may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of IGFAM and immunoglobulin superfamily proteins and Ig domain-containing proteins such as antibody heavy and light chains. In addition, the expression of IGFAM is closely associated with proliferating tissues, cancerous tissue and with hematopoiesis, inflammation, and other processes mediated by the immune system. Therefore, IGFAM appears to play a role in cancer, immune system disorders, and infections. In the treatment of disorders associated with increased IGFAM expression or activity, it is desirable to decrease the expression or activity of IGFAM. In the treatment of disorders associated with decreased IGFAM expression or activity, it is desirable to increase the expression or activity of IGFAM.

Therefore, in one embodiment, IGFAM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IGFAM. Examples of such disorders include, but are not limited to, a cancer such as

adenocarcinoma, melanoma, sarcoma, teratocarcinoma, and in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a disorder of the immune system such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other mycosis-causing fungal agent; and an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematode such as trichinella, intestinal nematode such as ascaris, lymphatic filarial nematode, trematode such as schistosoma, and cestrode such as tapeworm.

In another embodiment, a vector capable of expressing IGFAM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased

expression or activity of IGFAM including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified IGFAM in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IGFAM including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of IGFAM may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IGFAM including, but not limited to, those listed above.

In a further embodiment, an antagonist of IGFAM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of IGFAM. Examples of such disorders include, but are not limited to, those cancer, immune system disorders, and infections described above. In one aspect, an antibody which specifically binds IGFAM may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express IGFAM.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding IGFAM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of IGFAM including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of IGFAM may be produced using methods which are generally known in the art. In particular, purified IGFAM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind IGFAM. Antibodies to IGFAM may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with IGFAM or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to

increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium paryum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to IGFAM have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of IGFAM amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to IGFAM may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce IGFAM-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for IGFAM may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and

easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between IGFAM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering IGFAM epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for IGFAM. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of IGFAM-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple IGFAM epitopes, represents the average affinity, or avidity, of the antibodies for IGFAM. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular IGFAM epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the IGFAM-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of IGFAM, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of IGFAM-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding IGFAM, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding IGFAM may be used in situations in which it would be

desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding IGFAM. Thus, complementary molecules or fragments may be used to modulate IGFAM activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding IGFAM.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding IGFAM. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding IGFAM can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding IGFAM. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding IGFAM. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding IGFAM.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding IGFAM. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of IGFAM,

antibodies to IGFAM, and mimetics, agonists, antagonists, or inhibitors of IGFAM. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol.

Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of IGFAM, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of

administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example IGFAM or fragments thereof, antibodies of IGFAM, and agonists, antagonists or inhibitors of IGFAM, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind IGFAM may be used for the diagnosis of disorders characterized by expression of IGFAM, or in assays to monitor patients being treated with IGFAM or agonists, antagonists, or inhibitors of IGFAM. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for IGFAM include methods which utilize the antibody and a label to detect

IGFAM in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring IGFAM, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of IGFAM expression. Normal or standard values for IGFAM expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to IGFAM under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of IGFAM expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding IGFAM may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of IGFAM may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of IGFAM, and to monitor regulation of IGFAM levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding IGFAM or closely related molecules may be used to identify nucleic acid sequences which encode IGFAM. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding IGFAM, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the IGFAM encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:20-38 or from genomic sequences including promoters, enhancers, and introns of the IGFAM gene.

Means for producing specific hybridization probes for DNAs encoding IGFAM include the cloning of polynucleotide sequences encoding IGFAM or IGFAM derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a

variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding IGFAM may be used for the diagnosis of disorders associated with expression of IGFAM. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, melanoma, sarcoma, teratocarcinoma, and in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a disorder of the immune system such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other mycosis-causing fungal agent; and an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas,

tissue nematode such as trichinella, intestinal nematode such as ascaris, lymphatic filarial nematode, trematode such as schistosoma, and cestrode such as tapeworm. The polynucleotide sequences encoding IGFAM may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered IGFAM expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding IGFAM may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding IGFAM may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding IGFAM in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of IGFAM, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding IGFAM, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals

to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding IGFAM may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding IGFAM, or a fragment of a polynucleotide complementary to the polynucleotide encoding IGFAM, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of IGFAM include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding IGFAM may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding IGFAM on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, IGFAM, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between IGFAM and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with IGFAM, or fragments thereof, and washed. Bound IGFAM is then detected by methods well known in the art. Purified IGFAM can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing

antibodies capable of binding IGFAM specifically compete with a test compound for binding IGFAM. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with IGFAM.

In additional embodiments, the nucleotide sequences which encode IGFAM may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0643 P, filed November 19, 1998], U.S. Ser. No. 60/113,635, and U.S. Ser. No. 60/128,194, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA

libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene

families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:20-38. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding IGFAM occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of IGFAM Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:20-38 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this

fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on

antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:20-38 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:20-38 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the IGFAM-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring IGFAM. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of IGFAM. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the IGFAM-encoding transcript.

IX. Expression of IGFAM

Expression and purification of IGFAM is achieved using bacterial or virus-based expression systems. For expression of IGFAM in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid

promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express IGFAM upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of IGFAM in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding IGFAM by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, IGFAM is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from IGFAM at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified IGFAM obtained by these methods can be used directly in the following activity assay.

X. Demonstration of IGFAM Activity

An assay for IGFAM activity measures the ability of IGFAM to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E. S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.) IGFAM is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled IGFAM. IGFAM-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable IGFAM-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable IGFAM-antigen complex is plotted against the serum concentration. For various serum

concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable IGFAM-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable IGFAM-antigen complex is a measure of IGFAM activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for IGFAM activity measures the expression of IGFAM on the cell surface. cDNA encoding IGFAM is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et.al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using IGFAM-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of IGFAM expressed on the cell surface.

XI. Functional Assays

IGFAM function is assessed by expressing the sequences encoding IGFAM at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of IGFAM on gene expression can be assessed using highly purified

populations of cells transfected with sequences encoding IGFAM and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding IGFAM and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of IGFAM Specific Antibodies

IGFAM substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the IGFAM amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-IGFAM activity by, for example, binding the peptide or IGFAM to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring IGFAM Using Specific Antibodies

Naturally occurring or recombinant IGFAM is substantially purified by immunoaffinity chromatography using antibodies specific for IGFAM. An immunoaffinity column is constructed by covalently coupling anti-IGFAM antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing IGFAM are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of IGFAM (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/IGFAM binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and IGFAM is collected.

XIV. Identification of Molecules Which Interact with IGFAM

IGFAM, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled IGFAM, washed, and any wells with labeled IGFAM complex are assayed. Data obtained using different concentrations of IGFAM are used to calculate values for the number, affinity, and association of IGFAM with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	20	079785	SYNORAB01	079785H1 (SYNORAB01), 079785R6 (SYNORAB01), 930708R1 (CERVNOT01)
2	21	2469025	THYRNOT08	731672R1 (LUNGNOT03), 1241709R6 (LUNGNOT03), 1535284H1 (SPLNNOT04), 2242944T6 (PANCTUT02), 2354183H1 (LUNGNOT20), 2469025H1 (THYRNOT08), 3370185H1 (CONNTUT04), 3508954H1 (CONCNOT01), 3882135F6 (SPLNNOT11), 5021181H1 (OVARNON03)
3	22	2906265	THYMNOT05	2684470T6 (LUNGNOT23), 2906265H1 (THYMNOT05), 2906265T6 (THYMNOT05), SBQA01965D1
4	23	788975	PROSTUT03	334315R6 (EOSIHET02), 788975H1 (PROSTUT03), 2995078H1 (OVARTUT07), 3210491T6 (BLADNOT08), SAJA03793F1, SAJA00460R1
5	24	1407148	LATRTUT02	1407148H1 (LATRTUT02), 1666491T6 (BMARNOT03), 3339721F6 (SPLNNOT10)
6	25	1870848	SKINBIT01	1600994F6 (BLADNOT03), 1870848F6 (SKINBIT01), 1870848H1 (SKINBIT01), 1870848T3 (SKINBIT01), 3586454H1 (293TF4T01), SBGA04741F1, SBGA05485F1
7	26	1888468	BLADTUT07	1888468H1 (BLADTUT07), 2848829H1 (BRSTTUT13), 3336345F6 (SPLNNOT10)
8	27	2770104	COLANOT02	034758X22 (THP1NOB01), 034758X34R1 (THP1NOB01), 169537X14 (BMARNOR02), 1526466F1 (UCMCL5T01), 2770104H1 (COLANOT02)
9	28	2851053	BRSTTUT13	540284R1 (LNODNOT02), 689572R6 (LUNGTUT02), 2851053H1 (BRSTTUT13)

Table 1 cont.

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
10	29	3238787	COLAUCT01	540284F1 (LNODNOT02), 792219T1 (PROSTUT03), 2179956H1 (SININOT01), 2630239T6 (COLNTUT15), 3238787H1 (COLAUCT01)
11	30	3559548	LUNGNOT31	701560H1 (SYNORAT03), 3336345F6 (SPLNNOT10), 3559548H1 (LUNGNOT31)
12	31	3872741	BMARNOT03	2354696F6 (LUNGNOT20), 2998376F6 (OVARTUT07), 3872741H1 (BMARNOT03)
13	32	3981428	LUNGTUT08	930366T1 (CERVNOT01), 3981428F6 (LUNGTUT08), 3981428H1 (LUNGTUT08)
14	33	4635039	GBLADIT02	2849560F6 (BRSTTUT13), 4635039H1 (GBLADIT02)
15	34	3240710	COLAUCT01	1820516T6 (GBLATUT01), 3240710H1 (COLAUCT01), 3344660F6 (SPLNNOT09)
16	35	4945813	SINTNOT25	1435848F6 (PANCNOT08), 1822478F6 (GBLATUT01), 4945813H1 (SINTNOT25), SBJA01188F1, SBPA02549D1
17	36	4948957	SINTNOT25	2183186H1 (SININOT01), 4948957H1 (SINTNOT25), SAWA00061F1
18	37	4949649	SINTNOT25	232824F1 (SINTNOT02), 991833T1 (COLNNOT11), 2994953H1 (OVARTUT07), 3860911T6 (LNODNOT03), 4949649H1 (SINTNOT25)
19	39	5500302	BRABDIR01	036351X10, 036351X6 and 036351X9 (HUVENOB01), 777804R1 (COLNNOT05), 810532R1 (LUNGNOT04), 2538351H1 (BONRTUT01), 5500302H1 (BRABDIR01)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
1	237	S36 S89 T125 S185 T187 S205 T44		Ig domains G38-Q112 S150-V219 D193-E236 Ig signatures S154-A176 Y215-F232 Signal peptide M1-C22	κ L-chain (g2765423) (Homo sapiens)	BLAST MOTIFS PFAM PROFILESCAN BLOCKS HMM SPSCAN
2	537	S215 T117 T206 S246 S511 S93 T104 S110 S256 T284 T288 Y529	N83 N378 N469 N520	Ig domains S41-R128 N174-V239 R336-L404 E440-V510 T486-M536 Ig signatures Y235-H241 W444-R466 Y506-R523 Signal peptide M1-S26	δ H-chain (g495871) (Homo sapiens)	BLAST MOTIFS PFAM PROFILESCAN BLOCKS SPSCAN

Table 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
3	311	S72 T101 T132 S151 S238 S28 T32 S81 T158 S211	N204	Ig domains G35-S113 T158-S211 Ig signatures T162-K184 F228-Q245 Signal peptide M1-A21 Transmembrane domain Y283-M303	T cell receptor β chain (g3002935) (Homo sapiens)	BLAST MOTIFS PFAM BLOCKS HMM SPSCAN
4	194	S46 T70 S105 T119 T2 S43	N79 N94 N103 N147	Signal peptide M1-A30	MC51L-53L-54L homolog (secreted glycoprotein) (g5231020) (Molluscum contagiosum virus)	BLAST MOTIFS
5	236	S36 T124 S184 T186 S204 T44		Ig domains G38-Q112 S149-V218 D192-E235 Signal peptide M1-C22	Ig κ light chain (g2765423) (Homo sapiens)	BLAST MOTIFS PROFILESCAN BLOCKS PFAM SPSCAN HMM

Table 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
6	310	S110 T106 T296 S37 S99 S227 S281 Y77	N104 N192	Ig domains F46-V117 G153-A221 Signal peptide M1-G30 Transmembrane domain: Y238-T260	JAM: junctional adhesion molecule (g3462455) (Mus musculus)	BLAST MOTIFS PFAM SPSCAN HMM
7	148	S26 T106 T4 S36 S40 S142		Ig domain G34-R117 Signal peptide M1-S19	Ig heavy chain variable region (g3170981) (Homo sapiens)	BLAST MOTIFS PFAM SPSCAN HMM
8	310	T11 S303 T137 T236 S300	N183	Ig domains: G86-Y144 G181-S241 Transmembrane domains: E275-F295	ILT1c: myeloid- and lymphoid-specific Ig-like receptor (g1907323) (Homo sapiens)	BLAST MOTIFS PFAM HMM
9	236	S36 S89 T124 S184 T186 S204 T44 Y54		Ig domains G38-Q112 S149-V218 D192-E235 Signal peptide M1-C22	Ig κ light chain (g2765423) (Homo sapiens)	BLAST MOTIFS PROFILESCAN BLOCKS PFAM SPSCAN HMM

Table 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
10	237	S36 S89 T125 S185 T187 S205 T44		Ig domains G38-Q112 S150-V219 D193-E236 Signal peptide M1-C22	Ig κ light chain (g2765423) (Homo sapiens)	BLAST MOTIFS PROFILESCAN BLOCKS PFAM SPSCAN HMM
11	148	S26 T106 S36 S40 S142		Ig domain G34-R117 Signal peptide M1-S19	Ig heavy chain variable region (g3170891) (Homo sapiens)	BLAST MOTIFS PFAM SPSCAN HMM
12	236	S36 S74 T124 S184 T186 S204		Ig domains G38-H112 S149-V218 D192-E235 Signal peptide M1-C22	Ig κ light chain (g2765423) (Homo sapiens)	BLAST MOTIFS PROFILESCAN BLOCKS PFAM SPSCAN HMM
13	237	S36 S89 T125 S185 T187 S205 T44		Ig domains G38-Q112 S150-V219 D193-E236 Signal peptide M1-C22	Ig κ light chain (g2765423) (Homo sapiens)	BLAST MOTIFS PROFILESCAN BLOCKS PFAM SPSCAN HMM

Table 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
14	219	S26 T47 T157 T181 S183 T4 S36 S40 S146 T190		Ig domain G34-R117 Signal peptide M1-S19	Ig heavy chain variable region (g386807) (Homo sapiens)	BLAST MOTIFS PFAM BLOCKS SPSCAN HMM
15	241	S18 S92 S101 T129 S189 T191 S209 S42 S77 T97		Ig Domain: G36-R115 S154-V223 Signal Peptide: M1-A20	chimeric monoclonal TSH ab, κ chain (g413074) (synthetic construct)	BLAST PFAM, MOTIFS, BLOCKS, SPSCAN, ProfileScan
16	507	T114 T170 T194 S410 S435 S90 T101 S107 S159 T203 S320 T379 T479	N298 N494	Ig Domain: S41-T125 N397-V469 Signal Peptide: M1-S26	SNC73 protein, down regulated protein in colorectal cancer (g3201900) (Homo sapiens)	BLAST PFAM, MOTIFS, BLOCKS, SPSCAN, ProfileScan
17	244	S38 S82 S97 T132 S192 T194 S212 T46		Ig Domain: G40-Q120 S157-V226 Signal Peptide: M1-A20	anti-Entamoeba histolytica Ig κ light chain (g5360673) (Homo sapiens)	BLAST PFAM, MOTIFS, BLOCKS, SPSCAN, ProfileScan

Table 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
18	240	S20 S149 S228 S78 S215		Ig Domain: G39-L116 A155-V223 Signal Peptide: M1-S24	Ig λ light chain (g2765427) (Homo sapiens)	BLAST PFAM, MOTIFS, BLOCKS, SPSCAN, ProfileScan
19	398	T88 T107 T152 S199 S74 T194 S286 Y216	N171 N197 N229 N284 N296 N317 N382	Ig Domain: G45-A118 G252-A318 Signal Peptide: M1-A28 Transmembrane L340-Y360	Glycoprotein 56 (g1806278) (Rattus norvegicus)	BLAST PFAM, MOTIFS, BLOCKS, SPSCAN, HMMER, ProfileScan

Table 3

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Diseases, Disorders, or Conditions (Fraction of Total)	Vector
20	356-400	Reproductive (0.285) Gastrointestinal (0.261) Cardiovascular (0.116)	Cancer (0.598) Inflammation (0.269)	PBLUESCRIPT
21	110-154	Gastrointestinal (0.308) Reproductive (0.215) Hematopoietic/Immune (0.159)	Cancer (0.514) Inflammation (0.411)	pINCY
22	346-393	Hematopoietic/Immune (0.274) Gastrointestinal (0.233) Reproductive (0.178)	Inflammation (0.411) Cancer (0.356) Fetal (0.110)	pINCY
23	659-688	Hematopoietic/Immune (0.407) Nervous (0.148) Cardiovascular (0.111)	Inflammation (0.556) Cancer (0.333) Fetal (0.074)	PSPORT1
24	209-250	Gastrointestinal (0.270) Reproductive (0.266) Hematopoietic/Immune (0.129)	Cancer (0.571) Inflammation (0.296) Trauma (0.086)	pINCY
25	494-538	Nervous (0.412) Reproductive (0.265) Cardiovascular (0.088)	Cancer (0.471) Inflammation (0.206) Fetal (0.176)	pINCY
26	263-307	Gastrointestinal (0.324) Reproductive (0.250) Hematopoietic/Immune (0.120)	Cancer (0.546) Inflammation (0.370) Trauma (0.065)	pINCY
27	327-371 597-641 1029-1073	Hematopoietic/Immune (0.562) Gastrointestinal (0.188) Musculoskeletal (0.125)	Inflammation (0.625) Cancer (0.250) Fetal (0.250)	PINCY

Table 3 cont.

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Diseases, Disorders, or Conditions (Fraction of Total)	Vector
28	149-232	Reproductive (0.280) Gastrointestinal (0.259) Hematopoietic/Immune (0.126)	Cancer (0.582) Inflammation (0.285) Trauma (0.084)	pINCY
29	154-183	Reproductive (0.274) Gastrointestinal (0.259) Cardiovascular (0.124)	Cancer (0.586) Inflammation (0.263) Trauma (0.094)	pINCY
30	245-292	Gastrointestinal (0.316) Reproductive (0.194) Cardiovascular (0.143)	Cancer (0.571) Inflammation (0.337) Trauma (0.071)	pINCY
31	152-187	Gastrointestinal (0.281) Reproductive (0.255) Hematopoietic/Immune (0.128)	Cancer (0.570) Inflammation (0.294) Trauma (0.085)	pINCY
32	360-389	Reproductive (0.281) Gastrointestinal (0.258) Cardiovascular (0.123)	Cancer (0.585) Inflammation (0.269) Trauma (0.092)	pINCY
33	380-430	Gastrointestinal (0.306) Reproductive (0.218) Cardiovascular (0.137)	Cancer (0.565) Inflammation (0.339) Trauma (0.081)	pINCY
34	70-114 355-396	Gastrointestinal (0.269) Reproductive (0.269) Hematopoietic/Immune (0.130) Cardiovascular (0.121)	Cancer (0.561) Inflammation (0.305)	pINCY
35	80-121 177-466	Reproductive (0.284) Gastrointestinal (0.276) Cardiovascular (0.133) Hematopoietic/Immune (0.116)	Cancer (0.591) Inflammation (0.307)	pINCY

Table 3 cont.

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Diseases, Disorders, or Conditions (Fraction of Total)	Vector
36	1-32 (5' UTR, possible promoter) 105-149 393-434	Gastrointestinal (0.272) Reproductive (0.263) Hematopoietic/Immune (0.129) Cardiovascular (0.121)	Cancer (0.567) Inflammation (0.299)	pINCY
37	370-414	Gastrointestinal (0.303) Reproductive (0.244) Cardiovascular (0.129) Hematopoietic/Immune (0.119)	Cancer (0.562) Inflammation (0.308)	pINCY
38	2-169 (5'-UTR, possible promoter) 1250-1363	Reproductive (0.245) Nervous (0.214) Gastrointestinal (0.153) Cardiovascular (0.133)	Cancer (0.408) Inflammation (0.255)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
20	SYNORAB01	The SYNORAB01 cDNA library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.
21	THYRNOT08	The THYRNOT08 cDNA library was constructed using RNA isolated from diseased left thyroid tissue removed from a 13-year-old Caucasian female during a complete thyroidectomy. Pathology indicated lymphocytic thyroiditis.
22	THYMNOT05	The THYMNOT05 cDNA library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during thymectomy and closure of a patent ductus arteriosus. Patient history included cardiac catheterization.
23	PROSTUT03	The library was constructed using RNA isolated from prostate tumor tissue removed from a 67-year-old Caucasian male during radical prostatectomy and lymph node biopsy. Pathology indicated adenocarcinoma Gleason grade 3+3. Adenofibromatous hyperplasia was present. Patient history included coronary artery disease, stomach ulcer, and osteoarthritis. Family history included congestive heart failure.
24	LATRTUT02	The library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
25	SKINBIT01	The library was constructed using RNA isolated from diseased skin tissue of the left lower leg. Patient history included erythema nodosum of the left lower leg.

Table 4 cont.

Nucleotide SEQ ID NO:	Library	Library Description
26	BLADTUT07	The library was constructed using RNA isolated from bladder tumor tissue removed from the anterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated a grade 3 transitional cell carcinoma in the left lateral bladder. Patient history included angina, emphysema, and tobacco use. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
27	COLANOT02	The library was constructed using RNA isolated from diseased ascending colon tissue removed from a 25-year-old Caucasian female during a multiple segmental resection of the large bowel. Pathology indicated moderately to severely active chronic ulcerative colitis, involving the entire colectomy specimen and sparing 2 cm of the attached ileum. Grossly, the specimen showed continuous involvement from the rectum proximally; marked mucosal atrophy and no skip areas were identified. Microscopically, the specimen showed dense, predominantly mucosal inflammation and crypt abscesses. Patient history included benign large bowel neoplasm. Previous surgeries included a polypectomy.
28	BRSTTUT13	The library was constructed using RNA isolated from breast tumor tissue removed from the right breast of a 46-year-old Caucasian female during a unilateral extended simple mastectomy with breast reconstruction. Pathology indicated an invasive grade 3 adenocarcinoma, ductal type with apocrine features and greater than 50% intraductal component. Patient history included breast cancer.
29	COLAUCT01	The library was constructed using RNA isolated from diseased ascending colon tissue removed from a 74-year-old Caucasian male during a multiple-segment large bowel excision with temporary ileostomy. Pathology indicated inflammatory bowel disease most consistent with chronic ulcerative colitis, characterized by severe acute and chronic mucosal inflammation with erythema, ulceration, and pseudopolyp formation involving the entire colon and rectum. The sigmoid colon had an area of mild stricture formation. One diverticulum with diverticulitis was identified near this zone.

Table 4 cont.

Nucleotide SEQ ID NO:	Library	Library Description
30	LUNGNOT31	The library was constructed using RNA isolated from right middle lobe lung tissue removed from a 63-year-old Caucasian male. Pathology for the associated tumor indicated grade 3 adenocarcinoma. Patient history included an abdominal aortic aneurysm, cardiac dysrhythmia, atherosclerotic coronary artery disease, hiatal hernia, chronic sinusitis, and lupus. Family history included acute myocardial infarction and atherosclerotic coronary artery disease.
31	BMARNOT03	The library was constructed using RNA isolated from the left tibial bone marrow tissue of a 16-year-old Caucasian male during a partial left tibial ostectomy with free skin graft. Patient history included an abnormality of the red blood cells. Previous surgeries included bone and bone marrow biopsy, and soft tissue excision. Family history included osteoarthritis.
32	LUNGTUT08	The library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma. Patient history included atherosclerotic coronary artery disease, acute myocardial infarction, rectal cancer, asymptomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, lung cancer, type II diabetes, atherosclerotic coronary artery disease, and acute myocardial infarction.
33	GBLADIT02	Library was constructed using RNA isolated from diseased gallbladder tissue removed from an 18-year-old Caucasian female during cholecystectomy and incidental appendectomy. Pathology indicated acute and chronic cholecystitis with cholelithiasis. The gallbladder contained multiple fragments of stony material. The appendix showed lymphoid hyperplasia. Previous surgeries included total splenectomy. Family history included benign hypertension, acute myocardial infarction, and atherosclerotic coronary artery disease.

Table 4 cont.

Nucleotide SEQ ID NO:	Library	Library Description
34	COLAUCT01	The COLAUCT01 library was constructed using RNA isolated from diseased ascending colon tissue removed from a 74-year-old Caucasian male during a multiple-segment large bowel excision with temporary ileostomy. Pathology indicated inflammatory bowel disease most consistent with chronic ulcerative colitis, characterized by severe acute and chronic mucosal inflammation with erythema, ulceration, and pseudopolyp formation involving the entire colon and rectum. The sigmoid colon had an area of mild stricture formation. One diverticulum with diverticulitis was identified near this zone.
35	SINTNOT25	The SINTNOT25 library was constructed using RNA isolated from small intestine tissue removed from a 13-year-old Caucasian male, who died from a gunshot wound to the head. Family history included diabetes.
36	SINTNOT25	The SINTNOT25 library was constructed using RNA isolated from small intestine tissue removed from a 13-year-old Caucasian male, who died from a gunshot wound to the head. Family history included diabetes.
37	SINTNOT25	The SINTNOT25 library was constructed using RNA isolated from small intestine tissue removed from a 13-year-old Caucasian male, who died from a gunshot wound to the head. Family history included diabetes.
38	BRABDIR01	The BRABDIR01 library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of IGFAM, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of IGFAM, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
YUE, Henry
TANG, Y. Tom
CORLEY, Neil C.
GUEGLER, Karl J.
GORGONE, Gina A.
BAUGHN, Mariah R.
LU, Dyung Aina M.
LAL, Preeti
HILLMAN, Jennifer L.
YANG, Junming

<120> IMMUNOGLOBULIN SUPERFAMILY PROTEINS

<130> PF-0643 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/195,853; unassigned; 60/113,635; 60/128,194
<151> 1998-11-19; 1998-11-19; 1998-12-22; 1999-04-07

<160> 38

<170> PERL Program

<210> 1
<211> 237
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 079785CD1

<400> 1

Met	Asp	Met	Arg	Val	Pro	Ala	Gln	Leu	Leu	Gly	Leu	Leu	Leu	Leu	
1				5					10						15
Trp	Leu	Arg	Gly	Ala	Arg	Cys	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	
			20						25						30
Ser	Ser	Leu	Ser	Ala	Ser	Val	Gly	Asp	Arg	Val	Thr	Ile	Thr	Cys	
			35						40						45
Arg	Ala	Gly	Gln	Ser	Ile	Ser	Ser	Tyr	Leu	Asn	Trp	Tyr	Gln	Gln	
			50						55						60
Lys	Pro	Gly	Lys	Ala	Pro	Lys	Leu	Leu	Ile	Tyr	Ala	Ala	Ser	Ser	
			65						70						75
Leu	Gln	Ser	Gly	Val	Pro	Ser	Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	
			80						85						90
Thr	Asp	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu	Gln	Pro	Glu	Asp	Phe	
			95						100						105
Ala	Thr	Tyr	Tyr	Cys	Gln	Gln	Ser	Tyr	Ser	Thr	Pro	Pro	Ile	Thr	
			110						115						120

```
Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys Arg Thr Val Ala Ala
                125                 130                 135
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser
            140                 145                 150
Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg
            155                 160                 165
Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
            170                 175                 180
Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
            185                 190                 195
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
            200                 205                 210
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
            215                 220                 225
Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
            230                 235
```

<210> 2
<211> 537
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2469025CD1

<400> 2

```
Met Asp Leu Leu His Lys Asn Met Lys His Leu Trp Phe Phe Leu
  1               5                  10                  15
Leu Leu Val Ala Ala Pro Arg Trp Val Leu Ser Gln Val Gln Leu
            20                  25                  30
Gln Gln Trp Gly Ala Gly Leu Leu Lys Pro Ser Glu Thr Leu Ser
            35                  40                  45
Leu Thr Cys Ala Val Tyr Gly Gly Ser Phe Ser Gly Tyr Tyr Leu
            50                  55                  60
Ser Gly Tyr Tyr Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly
            65                  70                  75
Leu Glu Trp Ile Gly Glu Ile Asn His Ser Gly Ser Thr Asn Tyr
            80                  85                  90
Asn Pro Ser Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser
            95                 100                 105
Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp
           110                 115                 120
Thr Ala Val Tyr Tyr Cys Ala Arg Gly Arg Ser Asp Ser Ser Gly
           125                 130                 135
Ser Pro Tyr Gly Leu Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr
           140                 145                 150
Val Ser Ser Ala Pro Thr Lys Ala Pro Asp Val Phe Pro Ile Ile
           155                 160                 165
Ser Gly Cys Arg His Pro Lys Asp Asn Ser Pro Val Val Leu Ala
           170                 175                 180
Cys Leu Ile Thr Gly Tyr His Pro Thr Ser Val Thr Val Thr Trp
           185                 190                 195
Tyr Met Gly Thr Gln Ser Gln Pro Gln Arg Thr Phe Pro Glu Ile
```

```
                200                 205                 210
Gln Arg Arg Asp Ser Tyr Tyr Met Thr Ser Ser Gln Leu Ser Thr
                215                 220                 225
Pro Leu Gln Gln Trp Arg Gln Gly Glu Tyr Lys Cys Val Val Gln
                230                 235                 240
His Thr Ala Ser Lys Ser Lys Lys Glu Ile Phe Arg Trp Pro Glu
                245                 250                 255
Ser Pro Lys Ala Gln Ala Ser Ser Val Pro Thr Ala Gln Pro Gln
                260                 265                 270
Ala Glu Gly Ser Leu Ala Lys Ala Thr Thr Ala Pro Ala Thr Thr
                275                 280                 285
Arg Asn Thr Gly Arg Gly Gly Glu Glu Lys Lys Lys Glu Lys Glu
                290                 295                 300
Lys Glu Glu Gln Glu Glu Arg Glu Thr Lys Thr Pro Glu Cys Pro
                305                 310                 315
Ser His Thr Gln Pro Leu Gly Val Tyr Leu Leu Thr Pro Ala Val
                320                 325                 330
Gln Asp Leu Trp Leu Arg Asp Lys Ala Thr Phe Thr Cys Phe Val
                335                 340                 345
Val Gly Ser Asp Leu Lys Asp Ala His Leu Thr Trp Glu Val Ala
                350                 355                 360
Gly Lys Val Pro Thr Gly Gly Val Glu Glu Gly Leu Leu Glu Arg
                365                 370                 375
His Ser Asn Gly Ser Gln Ser Gln His Ser Arg Leu Thr Leu Pro
                380                 385                 390
Arg Ser Leu Trp Asn Ala Gly Thr Ser Val Thr Cys Thr Leu Asn
                395                 400                 405
His Pro Ser Leu Pro Pro Gln Arg Leu Met Ala Leu Arg Glu Pro
                410                 415                 420
Ala Ala Gln Ala Pro Val Lys Leu Ser Leu Asn Leu Leu Ala Ser
                425                 430                 435
Ser Asp Pro Pro Glu Ala Ala Ser Trp Leu Leu Cys Glu Val Ser
                440                 445                 450
Gly Phe Ser Pro Pro Asn Ile Leu Leu Met Trp Leu Glu Asp Gln
                455                 460                 465
Arg Glu Val Asn Thr Ser Gly Phe Ala Pro Ala Arg Pro Pro Pro
                470                 475                 480
Gln Pro Gly Ser Thr Thr Phe Trp Ala Trp Ser Val Leu Arg Val
                485                 490                 495
Pro Ala Pro Pro Ser Pro Gln Pro Ala Thr Tyr Thr Cys Val Val
                500                 505                 510
Ser His Glu Asp Ser Arg Thr Leu Leu Asn Ala Ser Arg Ser Leu
                515                 520                 525
Glu Val Ser Tyr Val Thr Asp His Gly Pro Met Lys
                530                 535
```

<210> 3
<211> 311
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2906265CD1

<400> 3
Met Gly Thr Arg Leu Leu Phe Trp Val Ala Phe Cys Leu Leu Gly
1 5 10 15
Ala Asp His Thr Gly Ala Gly Val Ser Gln Ser Pro Ser Asn Lys
20 25 30
Val Thr Glu Lys Gly Lys Asp Val Glu Leu Arg Cys Asp Pro Ile
35 40 45
Ser Gly His Thr Ala Leu Tyr Trp Tyr Arg Gln Ser Leu Gly Gln
50 55 60
Gly Leu Glu Phe Leu Ile Tyr Phe Gln Gly Asn Ser Ala Pro Asp
65 70 75
Lys Ser Gly Leu Pro Ser Asp Arg Phe Ser Ala Glu Arg Thr Gly
80 85 90
Gly Ser Val Ser Thr Leu Thr Ile Gln Arg Thr Gln Gln Glu Asp
95 100 105
Ser Ala Val Tyr Leu Cys Ala Ser Ser Phe Leu Ala Gly Arg Gly
110 115 120
Asn Thr Ile Tyr Phe Gly Glu Gly Ser Trp Leu Thr Val Val Glu
125 130 135
Asp Leu Asn Lys Val Phe Pro Pro Glu Val Ala Val Phe Glu Pro
140 145 150
Ser Glu Ala Glu Ile Ser His Thr Gln Lys Ala Thr Leu Val Cys
155 160 165
Leu Ala Thr Gly Phe Phe Pro Asp His Val Glu Leu Ser Trp Trp
170 175 180
Val Asn Gly Lys Glu Val His Ser Gly Val Ser Thr Asp Pro Gln
185 190 195
Pro Leu Lys Glu Gln Pro Ala Leu Asn Asp Ser Arg Tyr Cys Leu
200 205 210
Ser Ser Arg Leu Arg Val Ser Ala Thr Phe Trp Gln Asn Pro Arg
215 220 225
Asn His Phe Arg Cys Gln Val Gln Phe Tyr Gly Leu Ser Glu Asn
230 235 240
Asp Glu Trp Thr Gln Asp Arg Ala Lys Pro Val Thr Gln Ile Val
245 250 255
Ser Ala Glu Ala Trp Gly Arg Ala Asp Cys Gly Phe Thr Ser Val
260 265 270
Ser Tyr Gln Gln Gly Val Leu Ser Ala Thr Ile Leu Tyr Glu Ile
275 280 285
Leu Leu Gly Lys Ala Thr Leu Tyr Ala Val Leu Val Ser Ala Leu
290 295 300
Val Leu Met Ala Met Val Lys Arg Lys Asp Phe
305 310

<210> 4
<211> 194
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 788975CD1

<400> 4

```
Met Thr Met Arg His Asn Trp Thr Pro Asp Leu Ser Pro Leu Trp
 1               5                  10                  15
Val Leu Leu Leu Cys Ala His Val Val Thr Leu Leu Val Arg Ala
            20                  25                  30
Thr Pro Val Ser Gln Thr Thr Thr Ala Ala Thr Ala Ser Val Arg
        35                  40                  45
Ser Thr Lys Asp Pro Cys Pro Ser Gln Pro Pro Val Phe Pro Ala
    50                  55                  60
Ala Lys Gln Cys Pro Ala Leu Glu Val Thr Trp Pro Glu Val Glu
65                  70                  75
Val Pro Leu Asn Gly Thr Leu Ser Leu Ser Cys Val Ala Cys Ser
                80                  85                  90
Arg Phe Pro Asn Phe Ser Ile Leu Tyr Trp Leu Gly Asn Gly Ser
            95                  100                 105
Phe Ile Glu His Leu Pro Gly Arg Leu Trp Glu Gly Ser Thr Ser
        110                 115                 120
Arg Glu Arg Gly Ser Thr Gly Thr Gln Leu Cys Lys Ala Leu Val
    125                 130                 135
Leu Glu Gln Leu Thr Pro Ala Leu His Ser Thr Asn Phe Ser Cys
140                 145                 150
Val Leu Val Asp Pro Glu Gln Val Val Gln Arg His Val Val Leu
155                 160                 165
Ala Gln Leu Trp Ala Gly Leu Arg Ala Thr Leu Pro Pro Thr Gln
                170                 175                 180
Glu Ala Leu Pro Ser Ser His Ser Ser Pro Gln Gln Gln Gly
            185                 190
```

<210> 5
<211> 236
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1407148CD1

```
<400> 5
Met Asp Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu
 1               5                  10                  15
Trp Leu Pro Gly Ala Arg Cys Asp Ile Gln Leu Thr Gln Ser Pro
            20                  25                  30
Ser Phe Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
        35                  40                  45
Arg Ala Ser Gln Leu Ile Ser Asn His Leu Ala Trp Tyr Gln Gln
    50                  55                  60
Lys Pro Gly Arg Ala Pro Lys Leu Leu Val His Ser Ala Ser Ile
65                  70                  75
Leu Gln Ser Gly Val Pro Leu Arg Phe Ser Gly Ser Gly Tyr Gly
                80                  85                  90
Thr Glu Phe Thr Leu Thr Val Ala Ser Leu Gln Pro Glu Asp Ser
            95                  100                 105
Ala Thr Tyr Tyr Cys Gln Gln Arg Asn Gly Tyr Pro Ile Thr Phe
        110                 115                 120
Gly Gln Gly Thr Arg Leu Glu Ile Lys Arg Thr Val Ala Ala Pro
```

```
                125                 130                 135
Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
                140                 145                 150
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu
                155                 160                 165
Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn
                170                 175                 180
Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr
                185                 190                 195
Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
                200                 205                 210
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
                215                 220                 225
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
                230                 235
```

<210> 6
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1870848CD1

<400> 6

```
Met Ala Leu Arg Arg Pro Pro Arg Leu Arg Leu Cys Ala Arg Leu
  1                 5                  10                  15
Pro Asp Phe Phe Leu Leu Leu Leu Phe Arg Gly Cys Leu Ile Gly
                 20                  25                  30
Ala Val Asn Leu Lys Ser Ser Asn Arg Thr Pro Val Val Gln Glu
                 35                  40                  45
Phe Glu Ser Val Glu Leu Ser Cys Ile Ile Thr Asp Ser Gln Thr
                 50                  55                  60
Ser Asp Pro Arg Ile Glu Trp Lys Lys Ile Gln Asp Glu Gln Thr
                 65                  70                  75
Thr Tyr Val Phe Phe Asp Asn Lys Ile Gln Gly Asp Leu Ala Gly
                 80                  85                  90
Arg Ala Glu Ile Leu Gly Lys Thr Ser Leu Lys Ile Trp Asn Val
                 95                 100                 105
Thr Arg Arg Asp Ser Ala Leu Tyr Arg Cys Glu Val Val Ala Arg
                110                 115                 120
Asn Asp Arg Lys Glu Ile Asp Glu Ile Val Ile Glu Leu Thr Val
                125                 130                 135
Gln Val Lys Pro Val Thr Pro Val Cys Arg Val Pro Lys Ala Val
                140                 145                 150
Pro Val Gly Lys Met Ala Thr Leu His Cys Gln Glu Ser Glu Gly
                155                 160                 165
His Pro Arg Pro His Tyr Ser Trp Tyr Arg Asn Asp Val Pro Leu
                170                 175                 180
Pro Thr Asp Ser Arg Ala Asn Pro Arg Phe Arg Asn Ser Ser Ser
                185                 190                 195
His Leu Asn Ser Glu Thr Gly Thr Leu Val Phe Thr Ala Val His
                200                 205                 210
```

```
Lys Asp Asp Ser Gly Gln Tyr Tyr Cys Ile Ala Ser Asn Asp Ala
                215                 220                 225
Gly Ser Ala Arg Cys Glu Glu Gln Glu Met Glu Val Tyr Asp Leu
                230                 235                 240
Asn Ile Gly Gly Ile Ile Gly Gly Val Leu Val Val Leu Ala Val
                245                 250                 255
Leu Ala Leu Ile Thr Leu Gly Ile Cys Cys Ala Tyr Arg Arg Gly
                260                 265                 270
Tyr Phe Ile Asn Asn Lys Gln Asp Gly Glu Ser Tyr Lys Asn Pro
                275                 280                 285
Gly Lys Pro Asp Gly Val Asn Tyr Ile Arg Thr Asp Glu Glu Gly
                290                 295                 300
Asp Phe Arg His Lys Ser Ser Phe Val Ile
                305                 310
```

<210> 7
<211> 148
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1888468CD1

<400> 7

```
Met Asp Trp Thr Trp Arg Ile Leu Phe Leu Val Ala Ala Ala Thr
  1                 5                  10                  15
Gly Ala His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val
                 20                  25                  30
Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly
                 35                  40                  45
Tyr Thr Phe Thr Gly Tyr Tyr Met His Trp Val Arg Gln Ala Pro
                 50                  55                  60
Gly Gln Gly Leu Glu Trp Met Gly Trp Ile Ser Pro Asn Asn Gly
                 65                  70                  75
Asp Thr Phe Tyr Ala His Arg Leu Gln Asp Arg Val Thr Leu Thr
                 80                  85                  90
Thr Asp Thr Ser Ala Thr Thr Gly Tyr Met Glu Leu Arg Ser Leu
                 95                 100                 105
Thr Ser Asp Asp Thr Ala Ile Tyr Tyr Cys Ala Arg Gly Asp Tyr
                110                 115                 120
Gly Asn Ser Leu Asp His Trp Gly Gln Gly Asn Leu Val Thr Val
                125                 130                 135
Ser Ser Ala Ser Pro Thr Ser Pro Lys Gly Leu Pro Ala
                140                 145
```

<210> 8
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID NO: 2770104CD1

<400> 8
Met Arg Arg Thr Gln Pro Leu Ser Val His Thr Gly Trp Glu Gly
1 5 10 15
Gly Glu Ala Ile Ser Leu Cys Val Ser Leu Ser Arg Gln His Arg
20 25 30
Gly Leu Ile His Pro Gln Ser Arg Ala Val Gly Gly Asp Ala Met
35 40 45
Thr Pro Ile Val Thr Val Leu Ile Cys Leu Gly Leu Ser Leu Gly
50 55 60
Pro Arg Thr His Val Gln Thr Gly Thr Ile Pro Lys Pro Thr Leu
65 70 75
Trp Ala Glu Pro Asp Ser Val Ile Thr Gln Gly Ser Pro Val Thr
80 85 90
Leu Ser Cys Gln Gly Ser Leu Glu Ala Gln Glu Tyr Arg Leu Tyr
95 100 105
Arg Glu Lys Lys Ser Ala Ser Trp Ile Thr Arg Ile Arg Pro Glu
110 115 120
Leu Val Lys Asn Gly Gln Phe His Ile Pro Ser Ile Thr Trp Glu
125 130 135
His Thr Gly Arg Tyr Gly Cys Gln Tyr Tyr Ser Arg Ala Arg Trp
140 145 150
Ser Glu Leu Ser Asp Pro Leu Val Ala Gly Asp Asp Arg Ser Tyr
155 160 165
Gln Asn Pro Thr Ser Gln Pro Ser Pro Gly Pro Val Val Thr Pro
170 175 180
Gly Lys Asn Val Thr Leu Leu Cys Gln Ser Arg Gly Gln Phe His
185 190 195
Thr Phe Leu Leu Thr Lys Glu Gly Ala Gly His Pro Pro Leu His
200 205 210
Leu Arg Ser Glu His Gln Ala Gln Gln Asn Gln Ala Glu Phe Arg
215 220 225
Met Gly Pro Val Thr Ser Ala His Val Gly Thr Tyr Arg Cys Tyr
230 235 240
Ser Ser Leu Ser Ser Asn Pro Tyr Leu Leu Ser Leu Pro Ser Asp
245 250 255
Pro Leu Glu Leu Val Val Ser Ala Ser Leu Gly Gln His Pro Gln
260 265 270
Asp Tyr Thr Val Glu Asn Leu Ile Arg Met Gly Val Ala Gly Leu
275 280 285
Val Leu Val Val Leu Gly Ile Leu Leu Phe Glu Ala Gln His Ser
290 295 300
Gln Arg Ser Leu Gln Asp Ala Ala Gly Arg
305 310

<210> 9
<211> 236
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2851053CD1

<400> 9
Met Asp Met Arg Val Leu Ala Gln Leu Leu Gly Leu Leu Leu Leu
1 5 10 15
Cys Phe Pro Gly Ala Arg Cys Asp Ile Gln Met Thr Gln Ser Pro
20 25 30
Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
35 40 45
Arg Ala Ser Gln Asp Ile Ser Asn Tyr Leu Ala Trp Phe Gln Gln
50 55 60
Lys Pro Gly Thr Ala Pro Lys Ser Leu Ile Tyr Asp Thr Ser Ser
65 70 75
Leu Gln Ser Gly Val Pro Ser Lys Phe Ser Gly Ser Gly Ser Gly
80 85 90
Thr Asp Phe Thr Leu Thr Ile Asn Ser Leu Gln Pro Glu Asp Phe
95 100 105
Ala Thr Tyr Tyr Cys Gln Gln His His Ser Tyr Pro Leu Thr Phe
110 115 120
Gly Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro
125 130 135
Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
140 145 150
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu
155 160 165
Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn
170 175 180
Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr
185 190 195
Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
200 205 210
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
215 220 225
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
230 235

<210> 10
<211> 237
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3238787CD1

<400> 10
Met Asp Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu
1 5 10 15
Trp Leu Arg Gly Ala Arg Cys Asp Ile Gln Met Thr Gln Ser Pro
20 25 30
Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
35 40 45
Arg Ala Ser Gln Ser Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln
50 55 60
Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Ala Ala Ser Ser
65 70 75

```
Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
                80                  85                  90
Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe
                95                 100                 105
Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro Pro Ile Thr
               110                 115                 120
Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys Arg Thr Val Ala Ala
               125                 130                 135
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser
               140                 145                 150
Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg
               155                 160                 165
Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
               170                 175                 180
Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
               185                 190                 195
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
               200                 205                 210
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
               215                 220                 225
Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
               230                 235
```

<210> 11
<211> 148
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3559548CD1

<400> 11

```
Met Asp Trp Thr Trp Ser Ile Leu Phe Leu Val Ala Ala Ala Thr
  1               5                  10                  15
Gly Ala His Ser Gln Val His Leu Val Gln Ser Gly Ala Glu Val
                20                  25                  30
Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly
                35                  40                  45
Tyr Thr Phe Thr Ser His Gly Ile Thr Trp Val Arg Gln Ala Pro
                50                  55                  60
Gly Gln Gly Leu Glu Trp Met Gly Trp Ile Ser Pro Asn Asn Gly
                65                  70                  75
Asp Thr Phe Tyr Ala His Arg Leu Gln Asp Arg Val Thr Leu Thr
                80                  85                  90
Thr Asp Thr Ser Ala Thr Thr Gly Tyr Met Glu Leu Arg Ser Leu
                95                 100                 105
Thr Ser Asp Asp Thr Ala Ile Tyr Tyr Cys Ala Arg Gly Asp Tyr
               110                 115                 120
Gly Asn Ser Leu Asp His Trp Gly Gln Gly Asn Leu Val Thr Val
               125                 130                 135
Ser Ser Ala Ser Pro Thr Ser Pro Lys Gly Leu Pro Ala
               140                 145
```

<210> 12
<211> 236
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3872741CD1

<400> 12
Met Asp Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu
1 5 10 15
Trp Leu Ser Gly Ala Arg Cys Asp Thr Gln Met Thr Gln Ser Pro
20 25 30
Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Leu Thr Ile Thr Cys
35 40 45
Gln Ala Ser Glu Asp Val Ile Lys Tyr Val Asn Trp Tyr Gln Gln
50 55 60
Lys Pro Arg Lys Ala Pro Lys Leu Leu Ile His Asp Ala Ser Asn
65 70 75
Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
80 85 90
Thr Leu Phe Thr Phe Thr Ile Ser Asn Leu Gln Pro Glu Asp Val
95 100 105
Ala Thr Tyr Tyr Cys Gln His Tyr Ala Ser His Pro Leu Thr Phe
110 115 120
Gly Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro
125 130 135
Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
140 145 150
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu
155 160 165
Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn
170 175 180
Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr
185 190 195
Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
200 205 210
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
215 220 225
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
230 235

<210> 13
<211> 237
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3981428CD1

<400> 13
Met Asp Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu

```
 1               5                  10                  15
Trp Leu Arg Gly Ala Arg Cys Asp Ile Gln Met Thr Gln Ser Pro
            20                  25                  30
Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Met Thr Cys
        35                  40                  45
Arg Ala Ser Gln Ser Ile Ser Thr Tyr Leu Asn Trp Tyr Gln Gln
    50                  55                  60
Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Ala Ala Ser Ser
65                  70                  75
Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
                80                  85                  90
Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe
            95                  100                 105
Ala Thr Tyr Tyr Cys Gln Gln Ser Phe Asn Thr His Met Tyr Thr
        110                 115                 120
Phe Gly Gln Gly Thr Arg Leu Glu Met Lys Arg Thr Val Ala Ala
    125                 130                 135
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser
140                 145                 150
Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg
                155                 160                 165
Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
            170                 175                 180
Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
        185                 190                 195
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
    200                 205                 210
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
215                 220                 225
Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
                230                 235
```

<210> 14
<211> 219
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4635039CD1

<400> 14

```
Met Asp Trp Thr Trp Arg Ile Leu Phe Leu Val Ala Ala Val Thr
 1               5                  10                  15
Gly Val His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val
            20                  25                  30
Arg Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly
        35                  40                  45
Tyr Thr Phe Ser Asp His Tyr Ile His Trp Val Arg Gln Ala Pro
    50                  55                  60
Gly Gln Gly Leu Glu Trp Met Gly Trp Ile Asn Pro Asn Ser Gly
65                  70                  75
Gly Ala Arg Tyr Ala Gln Gly Phe Gln Gly Leu Val Thr Met Thr
                80                  85                  90
```

```
Arg Asp Thr Ser Ile Ser Thr Ala Tyr Leu Glu Leu Arg Gly Leu
                95                  100                 105
Arg Ser Asp Gly Ser Ala Val Tyr Phe Cys Ala Arg Gln Thr Thr
            110                 115                 120
Ser Ser Pro Val Gly Asp Ala Phe Asp Ile Trp Gly Gln Gly Thr
            125                 130                 135
Met Val Thr Val Ser Ser Ala Ser Pro Thr Ser Pro Lys Val Phe
            140                 145                 150
Pro Leu Ser Leu Cys Ser Thr Gln Pro Asp Gly Asn Val Val Ile
            155                 160                 165
Ala Cys Leu Val Gln Gly Phe Phe Pro Gln Glu Pro Leu Ser Val
            170                 175                 180
Thr Trp Ser Glu Thr Asp Gln Gly Val Thr Ala Lys Lys Leu Pro
            185                 190                 195
Thr Gln Pro Gly Cys Leu Arg Gly Thr Val Asn His Glu Gln Pro
            200                 205                 210
Ala Asp Pro Ala Gly Gln Asn Ser Ala
            215
```

<210> 15
<211> 241
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3240710CD1

<400> 15

```
Met Arg Leu Pro Ala Gln Leu Leu Gly Leu Leu Met Leu Trp Ile
  1               5                  10                  15
Pro Gly Ser Ser Ala Asp Ile Val Leu Thr Gln Thr Pro Leu Ser
            20                  25                  30
Leu Ser Val Thr Pro Gly Gln Pro Ala Ser Ile Ser Cys Lys Ser
            35                  40                  45
Ser Glu Ser Leu Leu His Thr Asp Gly Lys Thr Tyr Leu His Trp
            50                  55                  60
Phe Val Gln Lys Ala Gly Gln Pro Pro Gln Val Leu Met Tyr Glu
            65                  70                  75
Val Ser Asn Arg Phe Ser Gly Val Pro Asp Arg Phe Ser Gly Ser
            80                  85                  90
Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile Ser Arg Val Glu Ala
            95                  100                 105
Glu Asp Val Arg Ile Tyr Tyr Cys Met Arg Thr Ile Gln Val Pro
            110                 115                 120
Pro Thr Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
            125                 130                 135
Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
            140                 145                 150
Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
            155                 160                 165
Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala
            170                 175                 180
Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser
```

185 190 195
Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys
200 205 210
Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His
215 220 225
Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu
230 235 240
Cys

<210> 16
<211> 507
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4945813CD1

<400> 16
Met Asp Leu Leu Cys Lys Asn Met Lys His Leu Trp Phe Phe Leu
1 5 10 15
Leu Leu Val Ala Ala Pro Arg Trp Val Leu Ser Gln Leu Gln Leu
20 25 30
Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu Thr Leu Ser
35 40 45
Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr Asn His
50 55 60
Tyr Trp Gly Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp
65 70 75
Ile Gly Ser Ile Phe Tyr Thr Gly Asn Ser Tyr Tyr Asn Pro Ser
80 85 90
Leu Lys Ser Arg Leu Ala Ile Ser Val Asp Thr Ser Lys Ser Gln
95 100 105
Leu Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val
110 115 120
Tyr Tyr Cys Ala Thr Val Pro Lys Thr Arg Ser Arg Pro Arg Gly
125 130 135
Tyr Thr Tyr Gly Pro Phe Asp Phe Trp Gly Gln Gly Thr Leu Val
140 145 150
Thr Val Ser Ser Ala Ser Pro Thr Ser Pro Lys Val Phe Pro Leu
155 160 165
Ser Leu Cys Ser Thr Gln Pro Asp Gly Asn Val Val Ile Ala Cys
170 175 180
Leu Val Gln Gly Phe Phe Pro Gln Glu Pro Leu Ser Val Thr Trp
185 190 195
Ser Glu Ser Gly Gln Gly Val Thr Ala Arg Asn Phe Pro Pro Ser
200 205 210
Gln Asp Ala Ser Gly Asp Leu Tyr Thr Thr Ser Ser Gln Leu Thr
215 220 225
Leu Pro Ala Thr Gln Cys Leu Ala Gly Lys Ser Val Thr Cys His
230 235 240
Val Lys His Tyr Thr Asn Pro Ser Gln Asp Val Thr Val Pro Cys
245 250 255
Pro Val Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Thr Pro Pro

```
                260                 265                 270
Thr Pro Ser Pro Ser Cys Cys His Pro Arg Leu Ser Leu His Arg
                275                 280                 285
Pro Ala Leu Glu Asp Leu Leu Leu Gly Ser Glu Ala Asn Leu Thr
                290                 295                 300
Cys Thr Leu Thr Gly Leu Arg Asp Ala Ser Gly Val Thr Phe Thr
                305                 310                 315
Trp Thr Pro Ser Ser Gly Lys Ser Ala Val Gln Gly Pro Pro Glu
                320                 325                 330
Arg Asp Leu Cys Gly Cys Tyr Ser Val Ser Ser Val Leu Pro Gly
                335                 340                 345
Cys Ala Glu Pro Trp Asn His Gly Lys Thr Phe Thr Cys Thr Ala
                350                 355                 360
Ala Tyr Pro Glu Ser Lys Thr Pro Leu Thr Ala Thr Leu Ser Lys
                365                 370                 375
Ser Gly Asn Thr Phe Arg Pro Glu Val His Leu Leu Pro Pro Pro
                380                 385                 390
Ser Glu Glu Leu Ala Leu Asn Glu Leu Val Thr Leu Thr Cys Leu
                395                 400                 405
Ala Arg Gly Phe Ser Pro Lys Asp Val Leu Val Arg Trp Leu Gln
                410                 415                 420
Gly Ser Gln Glu Leu Pro Arg Glu Lys Tyr Leu Thr Trp Ala Ser
                425                 430                 435
Arg Gln Glu Pro Ser Gln Gly Thr Thr Thr Phe Ala Val Thr Ser
                440                 445                 450
Ile Leu Arg Val Ala Ala Glu Asp Trp Lys Lys Gly Asp Thr Phe
                455                 460                 465
Ser Cys Met Val Gly His Glu Ala Leu Pro Leu Ala Phe Thr Gln
                470                 475                 480
Lys Thr Ile Asp Arg Leu Ala Gly Lys Pro Thr His Val Asn Val
                485                 490                 495
Ser Val Val Met Ala Glu Val Asp Gly Thr Cys Tyr
                500                 505
```

<210> 17
<211> 244
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4948957CD1

<400> 17

```
Met Val Leu Gln Thr Gln Val Phe Ile Ser Leu Leu Leu Trp Ile
  1               5                  10                  15
Ser Val Leu Thr Ala Gly Ala Tyr Gly Asp Ile Val Met Thr Gln
                20                  25                  30
Ser Pro Asp Ser Leu Ala Val Ser Leu Gly Glu Arg Ala Thr Ile
                35                  40                  45
Thr Cys Lys Ser Ser Gln Ser Val Phe Tyr Asn Ser Asn Asn Lys
                50                  55                  60
Asn Tyr Leu Val Trp Tyr Gln Gln Arg Pro Gly Gln Pro Pro Lys
                65                  70                  75
```

```
Met Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp
                80                  85                  90
Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile
                95                 100                 105
Ser Ser Leu Gln Ala Glu Asp Val Ala Leu Tyr Tyr Cys Gln Gln
               110                 115                 120
Tyr Phe Thr Thr Pro Tyr Thr Phe Gly Gln Gly Thr Arg Leu Glu
               125                 130                 135
Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro
               140                 145                 150
Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
               155                 160                 165
Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val
               170                 175                 180
Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu
               185                 190                 195
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr
               200                 205                 210
Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu
               215                 220                 225
Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn
               230                 235                 240
Arg Gly Glu Cys
```

<210> 18
<211> 240
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4949649CD1

```
<400> 18
Met Ser Val Pro Thr Met Ala Trp Met Met Leu Leu Leu Gly Leu
  1                 5                  10                  15
Leu Ala Tyr Gly Ser Gly Val Asp Ser Gln Thr Val Val Thr Gln
                20                  25                  30
Glu Pro Ser Leu Ser Val Ser Pro Gly Gly Thr Val Thr Leu Thr
                35                  40                  45
Cys Gly Leu Ala Ser Asp Ser Val Ser Thr Asn Phe Phe Pro Thr
                50                  55                  60
Trp Tyr Gln Gln Thr Pro Gly Gln Ala Pro Arg Thr Leu Ile Tyr
                65                  70                  75
Ser Thr Ser Thr Arg Ser Ser Gly Val Pro Asp Arg Phe Ser Gly
                80                  85                  90
Ser Ile Leu Gly Asn Lys Ala Ala Leu Thr Ile Thr Gly Ala Gln
                95                 100                 105
Ala Asp Asp Glu Ser Asp Tyr Tyr Cys Ala Leu Tyr Met Gly Ser
               110                 115                 120
Gly Ile Ser Val Phe Gly Gly Gly Thr Lys Val Thr Val Leu Gly
               125                 130                 135
Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser Ser
               140                 145                 150
```

```
Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser
                155                 160                 165
Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp Ser
            170                 175                 180
Ser Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro Ser Lys Gln
            185                 190                 195
Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu Thr Pro
            200                 205                 210
Glu Gln Trp Lys Ser His Arg Ser Tyr Ser Cys Gln Val Thr His
            215                 220                 225
Glu Gly Ser Thr Val Glu Lys Thr Val Ala Pro Thr Glu Cys Ser
            230                 235                 240
```

<210> 19
<211> 398
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 5500302CD1

<400> 19

```
Met Ser Gly Ser Ser Leu Pro Ser Ala Leu Ala Leu Ser Leu Leu
  1               5                  10                  15
Leu Val Ser Gly Ser Leu Leu Pro Gly Pro Gly Ala Ala Gln Asn
            20                  25                  30
Ala Gly Phe Val Lys Ser Pro Met Ser Glu Thr Lys Leu Thr Gly
            35                  40                  45
Asp Ala Phe Glu Leu Tyr Cys Asp Val Val Gly Ser Pro Thr Pro
            50                  55                  60
Glu Ile Gln Trp Trp Tyr Ala Glu Val Asn Arg Ala Glu Ser Phe
            65                  70                  75
Arg Gln Leu Trp Asp Gly Ala Arg Lys Arg Arg Val Thr Val Asn
            80                  85                  90
Thr Ala Tyr Gly Ser Asn Gly Val Ser Val Leu Arg Ile Thr Arg
            95                  100                 105
Leu Thr Leu Glu Asp Ser Gly Thr Tyr Glu Cys Arg Ala Ser Asn
            110                 115                 120
Asp Pro Lys Arg Asn Asp Leu Arg Gln Asn Pro Ser Ile Thr Trp
            125                 130                 135
Ile Arg Ala Gln Ala Thr Ile Ser Val Leu Gln Lys Pro Arg Ile
            140                 145                 150
Val Thr Ser Glu Glu Val Ile Ile Arg Asp Ser Pro Val Leu Pro
            155                 160                 165
Val Thr Leu Gln Cys Asn Leu Thr Ser Ser Ser His Thr Leu Thr
            170                 175                 180
Tyr Ser Tyr Trp Thr Lys Asn Gly Val Glu Leu Ser Ala Thr Arg
            185                 190                 195
Lys Asn Ala Ser Asn Met Glu Tyr Arg Ile Asn Lys Pro Arg Ala
            200                 205                 210
Glu Asp Ser Gly Glu Tyr His Cys Val Tyr His Phe Val Ser Ala
            215                 220                 225
Pro Lys Ala Asn Ala Thr Ile Glu Val Lys Ala Ala Pro Asp Ile
```

```
                230                 235                 240
Thr Gly His Lys Arg Ser Glu Asn Lys Asn Glu Gly Gln Asp Ala
                245                 250                 255
Thr Met Tyr Cys Lys Ser Val Gly Tyr Pro His Pro Asp Trp Ile
                260                 265                 270
Trp Arg Lys Lys Glu Asn Gly Met Pro Met Asp Ile Val Asn Thr
                275                 280                 285
Ser Gly Arg Phe Phe Ile Ile Asn Lys Glu Asn Tyr Thr Glu Leu
                290                 295                 300
Asn Ile Val Asn Leu Gln Ile Thr Glu Asp Pro Gly Glu Tyr Glu
                305                 310                 315
Cys Asn Ala Thr Asn Ala Ile Gly Ser Ala Ser Val Val Thr Val
                320                 325                 330
Leu Arg Val Arg Ser His Leu Ala Pro Leu Trp Pro Phe Leu Gly
                335                 340                 345
Ile Leu Ala Glu Ile Ile Ile Leu Val Val Ile Ile Val Val Tyr
                350                 355                 360
Glu Lys Arg Lys Arg Pro Asp Glu Val Pro Asp Asp Asp Glu Pro
                365                 370                 375
Ala Gly Pro Met Lys Thr Asn Ser Thr Asn Asn His Lys Asp Lys
                380                 385                 390
Asn Leu Arg Gln Arg Asn Thr Asn
                395
```

<210> 20
<211> 917
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 079785CB1

<400> 20

```
gtcaggacac agcatggaca tgagggtccc cgctcagctc ctggggctcc tgctactctg  60
gctccgaggt gccagatgtg acatccagat gacccagtct ccatcctccc tgtctgcatc 120
tgtaggagac agagtcacca tcacttgccg ggcaggtcag agcattagca gctatttaaa 180
ttggtatcag cagaaaccag ggaaagcccc taagctcctg atctatgctg catccagttt 240
gcaaagtggg gtcccatcaa ggttcagtgg cagtggatct gggacagatt tcactctcac 300
catcagcagt ctgcaacctg aagattttgc aacttactac tgtcaacaga gttacagtac 360
ccctccgatc accttcggcc aagggacacg actggagatt aaacgaactg tggctgcacc 420
atctgtcttc atcttcccgc catctgatga gcagttgaaa tctggaactg cctctgttgt 480
gtgcctgctg aataacttct atcccagaga ggccaaagta cagtggaagg tggataacgc 540
cctccaatcg ggtaactccc aggagagtgt cacagagcag gacagcaagg acagcaccta 600
cagcctcagc agcaccctga cgctgagcaa agcagactac gagaaacaca aagtctacgc 660
ctgcgaagtc acccatcagg gcctgagctc gcccgtcaca aagagcttca acaggggaga 720
gtgttagagg gagaagtgcc cccacctgct cctcagttcc agcctgaccc cctcccatcc 780
tttggcctct gacccttttt ccacagggga cctacccctа ttgcggtcct ccagctcatc 840
tttcacctca ccccccctcct cctccttggc tttaattatg ctaatgttgg aggagaatga 900
ataaataaag tgatcga                                                917
```

<210> 21
<211> 1746

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2469025CB1

<400> 21
ctgctcacat gggaaatact ttctgagagt catggacctc ctgcacaaga acatgaaaca 60
cctgtggttc ttcctcctcc tggtggcagc tcccagatgg gtcctgtccc aggtgcagct 120
acagcagtgg ggcgcaggac tgttgaagcc ttcggagacc ctgtccctca cctgcgctgt 180
ctatggtggg tccttcagtg gttactactt aagtggttac tactggagct ggatccgcca 240
gcccccaggg aaggggctgg agtggattgg ggaaatcaat catagtggaa gcaccaacta 300
caacccgtcc ctcaagagtc gagtcaccat atcagtagac acgtccaaga accagttctc 360
cctgaagctg agctctgtga ccgccgcgga cacggctgtg tattactgtg cgagaggcag 420
gagtgatagt agtggttccc catatggact tgactactgg ggccagggaa ccctggtcac 480
cgtctcctca gcacccacca aggctccgga tgtgttcccc atcatatcag ggtgcagaca 540
cccaaaggat aacagccctg tggtcctggc atgcttgata actgggtacc acccaacgtc 600
cgtgactgtc acctggtaca tggggacaca gagccagccc cagagaacct tccctgagat 660
acaaagacgg gacagctact acatgacaag cagccagctc tccacccccc tccagcagtg 720
gcgccaaggc gagtacaaat gcgtggtcca gcacaccgcc agcaagagta agaaggagat 780
cttccgctgg ccagagtctc caaaggcaca ggcctcctca gtgcccactg cacaacccca 840
agcagagggc agcctcgcca aggcaaccac agccccagcc accacccgta acacaggaag 900
aggaggagaa gagaagaaga aggagaagga gaaagaggaa caagaagaga gagagacaaa 960
gacaccagag tgtccgagcc acacccagcc tcttggcgtc tacctgctaa cccctgcagt 1020
gcaggacctg tggctccggg acaaagccac cttcacctgc ttcgtggtgg gcagtgacct 1080
gaaggatgct cacctgacct gggaggtggc tgggaaggtc cccacagggg gcgtggagga 1140
agggctgctg gagcggcaca gcaacggctc ccagagccag cacagccgtc tgaccctgcc 1200
caggtccttg tggaacgcgg ggacctccgt cacctgcaca ctgaaccatc ccagcctccc 1260
accccagagg ttgatggcgc tgagagaacc cgctgcgcag gcacccgtca agctttccct 1320
gaacctgctg gcctcgtctg accctcccga ggcggcctcg tggctcctgt gtgaggtgtc 1380
tggcttctcg ccccccaaca tcctcctgat gtggctggag gaccagcgtg aggtgaacac 1440
ttctgggttt gcccccgcac gccccctctc acagcccggg agcaccacgt tctgggcctg 1500
gagtgtgctg cgtgtcccag ccccgcccag ccctcagcca gccacctaca cgtgtgtggt 1560
cagccacgag gactcccgga ctctgctcaa cgccagccgg agcctagaag tcagctatgt 1620
aacagaccat ggccccatga aatgatcccg gaccagatcc gtccacaccc gccactcagc 1680
agctctggcc gagctcacag tacaaccaca ataaactctt gttgaatgaa ctctaaaaaa 1740
aaaaaa 1746

<210> 22
<211> 1160
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2906265CB1

<400> 22
gccatgggca ccaggctcct cttctgggtg gccttctgtc tcctgggggc agatcacaca 60
ggagctggag tctcccagtc ccccagtaac aaggtcacag agaagggaaa ggatgtagag 120
ctcaggtgtg atccaatttc aggtcatact gccctttact ggtaccgaca gagcctgggg 180
cagggcctgg agtttttaat ttacttccaa ggcaacagtg caccagacaa atcagggctg 240
cccagtgatc gcttctctgc agagaggact gggggatccg tctccactct gacgatccag 300

cgcacacagc aggaggactc ggccgtgtat ctctgtgcca gcagctttct tgcagggagg 360
ggaaacacca tatattttgg agagggaagt tggctcactg ttgtagagga cctgaacaag 420
gtgttcccac ccgaggtcgc tgtgtttgag ccatcagaag cagagatctc ccacacccaa 480
aaggccacac tggtgtgcct ggccacaggc ttcttccctg accacgtgga gctgagctgg 540
tgggtgaatg ggaaggaggt gcacagtggg gtcagcacgg acccgcagcc cctcaaggag 600
cagcccgccc tcaatgactc cagatactgc ctgagcagcc gcctgagggt ctcggccacc 660
ttctggcaga acccccgcaa ccacttccgc tgtcaagtcc agttctacgg gctctcggag 720
aatgacgagt ggacccagga tagggccaaa cccgtcaccc agatcgtcag cgccgaggcc 780
tggggtagag cagactgtgg ctttacctcg gtgtcctacc agcaaggggt cctgtctgcc 840
accatcctct atgagatcct gctagggaag gccaccctgt atgctgtgct ggtcagcgcc 900
cttgtgttga tggccatggt caagagaaag gatttctgaa ggcagccctg gaagtggagt 960
taggagcttc taacccgtca tggtttcaat acacattctt cttttgccag cgcttctgaa 1020
gagctgctct cacctctctg catcccaata gatatccccc tatgtgcatg cacacctgca 1080
cactcacggc tgaaatctcc ctaacccagg gggaccttag catgcctaag tgactaaacc 1140
aaataaaaat gttctggtca 1160

<210> 23
<211> 1356
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 788975CB1

<400> 23
gctgaagaga gcactgcctc cctgtgtgac tggagaagag gacgttgtca cagataaaga 60
gccaggctca ccagctcctg acgcatgcat catgaccatg agacacaact ggacaccaga 120
cctcagccct ttgtgggtcc tgctcctgtg tgcccacgtc gtcactctcc tggtcagagc 180
cacacctgtc tcgcagacca ccacagctgc cactgcctca gttagaagca caaaggaccc 240
ctgccctcc cagcccccag tgttcccagc agctaagcag tgtccagcat tggaagtgac 300
ctggccagag gtggaagtgc cactgaatgg aacgctgagc ttatcctgtg tggcctgcag 360
ccgcttcccc aacttcagca tcctctactg gctgggcaat ggttccttca ttgagcacct 420
cccaggccga ctgtgggagg ggagcaccag ccgggaacgt gggagcacag gtacgcagct 480
gtgcaaggcc ttggtgctgg agcagctgac ccctgccctg cacagcacca acttctcctg 540
tgtgctcgtg gaccctgaac aggttgtcca gcgtcacgtc gtcctggccc agctctgggc 600
tgggctgagg gcaaccttgc cccccaccca agaagccctg ccctccagcc acagcagtcc 660
acagcagcag ggttaagact cagcacaggg ccagcagcag cacaaccttg accagagctt 720
gggtcctacc tgtctacctg gagtgaacag tccctgactg cctgtaggct gcgtggatgc 780
gcaacacacc ccctccttct ctgctttggg tcccttctct caccaaattc aaactccatt 840
cccacctacc tagaaaatca cagcctcctt ataatgcctc ctcctcctgc cattctctct 900
ccacctatcc attagccttc ctaacgtcct actcctcaca ctgctctact gctcagaaac 960
caccaagact gttgatgcct tagccttgca ctccagggcc ctacctgcat ttcccacatg 1020
actttctgga agcctcccaa ctattcttgc ttttcccaga cagctcccac tcccatgtct 1080
ctgctcattt agtcccgtct tcctcaccgc cccagcaggg gaacgctcaa gcctggttga 1140
aatgctgcct cttcagtgaa gtcatcctct ttcagctctg gccgcattct gcagacttcc 1200
tatcttcgtg ctgtatgttt tttttttttcc cccttcactc taatggactg ttccagggaa 1260
gggatggggg cagcagctgc ttcggatcca cactgtatct gtgtcatccc cacatgggtc 1320
ctcataaagg attattcaat ggaggcaaaa aaaaaa 1356

<210> 24
<211> 916
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1407148CB1

<400> 24
gtcaggacac agcatggaca tgagggtccc cgctcagctc ctggggctcc tgctgctctg 60
gctcccaggt gccagatgtg acatccagtt gacccagtct ccatccttcc tgtctgcgtc 120
tgtgggagac agagtcacca tcacttgccg ggccagtcag ctcattagta atcatttagc 180
ctggtatcag caaaagccag ggagagcccc taaactcctg gtccatagtg catccattct 240
gcaaagtggg gtcccattaa gattcagcgg cagtggatac gggacagagt tcactctcac 300
agtcgccagc ctgcagcctg aggattccgc aacttattac tgtcaacagc gcaacggtta 360
tccgatcacc ttcggccaag ggacacgcct ggagattaaa cgaactgtgg ctgcaccatc 420
tgtcttcatc ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg 480
cctgctgaat aacttctatc ccagagaggc caaagtacag tggaaggtgg ataacgccct 540
ccaatcgggt aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag 600
cctcagcagc accctgacgc tgagcaaagc agactacgag aaacacaaag tctacgcctg 660
cgaagtcacc catcagggcc tgagctcgcc cgtcacaaag agcttcaaca ggggagagtg 720
ttagagggag aagtgccccc acctgctcct cagttccagc ctgacccccc cccatccttt 780
ggcctctgac ccttttcca caggggacct acccctattg cggtcctcca gctcatcttt 840
cacctcaccc ccctcctcct ccttggcttt aattatgcta atgttggagg agaatgaata 900
aataaagtga atcttc 916

<210> 25
<211> 1956
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1870848CB1

<400> 25
cggggactac aagccgcgcc gcgctgccgc tggcccctca gcaaccctcg acatggcgct 60
gaggcggcca ccgcgactcc ggctctgcgc tcggctgcct gacttcttcc tgctgctgct 120
tttcaggggc tgcctgatag gggctgtaaa tctcaaatcc agcaatcgaa ccccagtggt 180
acaggaattt gaaagtgtgg aactgtcttg catcattacg gattcgcaga caagtgaccc 240
caggatcgag tggaagaaaa ttcaagatga acaaaccaca tatgtgtttt ttgacaacaa 300
aattcaggga gacttggcgg gtcgtgcaga aatactgggg aagacatccc tgaagatctg 360
gaatgtgaca cggagagact cagcccttta tcgctgtgag gtcgttgctc gaaatgaccg 420
caaggaaatt gatgagattg tgatcgagtt aactgtgcaa gtgaagccag tgacccctgt 480
ctgtagagtg ccgaaggctg taccagtagg caagatggca acactgcact gccaggagag 540
tgagggccac ccccggcctc actacagctg gtatcgcaat gatgtaccac tgcccacgga 600
ttccagagcc aatcccagat ttcgcaattc ttcttcccac ttaaactctg aaacaggcac 660
tttggtgttc actgctgttc acaaggacga ctctgggcag tactactgca ttgcttccaa 720
tgacgcaggc tcagccaggt gtgaggagca ggagatggaa gtctatgacc tgaacattgg 780
cggaattatt gggggggttc tggttgtcct tgctgtactg gccctgatca cgttgggcat 840
ctgctgtgca tacagacgtg gctacttcat caacaataaa caggatggag aaagttacaa 900
gaacccaggg aaaccagatg gagttaacta catccgcact gacgaggagg gcgacttcag 960
acacaagtca tcgtttgtga tctgagaccc gcggtgtggc tgagagcgca cagagcgcac 1020
gtgcacatac ctctgctaga aactcctgtc aaggcagcga gagctgatgc actcggacag 1080
agctagacac tcattcagaa gcttttcgtt ttggccaaag ttgaccacta ctcttcttac 1140
tctaacaagc cacatgaata gaagaatttt cctcaagatg gacccggtaa atataaccac 1200

aaggaagcga aactgggtgc gttcactgag ttgggttcct aatctgtttc tggcctgatt 1260
cccgcatgaa tattagggtg atcttaaaga gtttgctcac gtaaacgccc gtgctgggcc 1320
ctgtgaagcc agcatgttca ccactggtcg ttcagcagcc acgacagcac catgtgagat 1380
ggcgaggtgg ctggacagca ccagcagcgc atcccggcgg gaacccagaa aaggcttctt 1440
acacagcagc cttacttcat cggcccacag acaccaccgc agtttcttct taaaggctct 1500
gctgatcggt gttgcagtgt ccattgtgga gaagcttttt ggatcagcat tttgtaaaaa 1560
caaccaaaat caggaaggta aattggttgc tggaagaggg atcttgcctg aggaaccctg 1620
cttgtccaac agggtgtcag gatttaagga aaaccttcgt cttaggctaa gtctgaaatg 1680
gtactgaaat atgcttttct atgggtcttg tttattttat aaaatttttac atctaaattt 1740
ttgctaagga tgtattttga ttattgaaaa gaaaatttct atttaaactg taaatatatt 1800
gtcatacaat gttaaataac ctatttttt taaaaaagtt caacttaagg tagaagttcc 1860
aagctactag tgataaaatg ggaaatatca ataattaaga gtattttacc caaggaaatc 1920
tcctcagggg agtttactgt gatggtcttt tcccgc 1956

<210> 26
<211> 589
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 1888468CB1

<400> 26
cagcaaccac atctgtcctc tagagaatcc cctgagagct ccgttcctca ccatggactg 60
gacctggagg atcctcttct tggtggcagc agccacagga gcccactccc aggtgcagct 120
ggtgcagtct ggggctgagg tgaagaagcc tggggcctca gtgaaggtct cctgcaaggc 180
ttctggatac accttcaccg gctactatat gcactgggtg cgacaggccc ctggacaagg 240
gcttgagtgg atggggtgga tcagccctaa caatggtgac acattctatg cacacagact 300
ccaggacaga gtcaccttga ccacagacac atcggcgacc acaggctaca tggagctgag 360
gagcctgaca tctgacgaca cggccatata ttattgcgcg agaggggact acggtaactc 420
tcttgaccac tggggccagg gaaacctggt cactgtctcc tcagcatccc cgaccagccc 480
caaaggtctt cccgcttgag cctctgcaag cacccagccc agatggggaa ccgtggtcaa 540
tcgcctgccc tgggtccagg ggctttcttc cccaagggag ccaattaat 589

<210> 27
<211> 1388
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2770104CB1

<400> 27
caggaaggag cagacacaag aagcaccagt tctatttgct gctacatccc ggctctcggt 60
gagacgagga gaaaccagac agacagtggc tgggggtcag gaaagacccc attacagtct 120
gaaatgtctg cagagggccc agtgcatgcn cccacctcag ctctaaaaga atgagagtca 180
ggctcctggg agggcagttc cgcttcttgt gtggctgcag atgacaacac cccatgagaa 240
ggacccagcc tctgagtgtc cacacagggt gggaaggagg ggaggctatt tctctctgtg 300
tgtctctgtc ccgccagcac cgagggctca tccatccgca gagcagggca gtgggaggag 360
acgccatgac ccccatcgtc acagtcctga tctgtctcgg gctgagtctg ggccccagga 420
cccacgtgca gacagggacc atccccaagc ccaccctgtg ggctgagcca gactctgtga 480

tcacccaggg gagtcccgtc accctcagtt gtcaggggag ccttgaagcc caggagtacc 540
gtctatatag ggagaaaaaa tcagcatctt ggattacacg gatacgacca gagcttgtga 600
agaacggcca gttccacatc ccatccatca cctgggaaca cacagggcga tatggctgtc 660
agtattacag ccgcgctcgg tggtctgagc tcagtgaccc cctggttgct ggtgatgaca 720
ggagctacca aaacccaacc tctcagccca gcccaggccc tgtggtgacc ccaggaaaga 780
acgtgaccct gctgtgtcag tcacgggggc agttccacac tttccttctg accaaggagg 840
gggcaggcca tcccccactg catctgagat cagagcacca agctcagcag aaccaggctg 900
aattccgcat gggtcctgtg acctcagccc acgtggggac ctacagatgc tacagctcac 960
tcagctccaa cccctacctg ctgtctctcc ccagtgaccc cctggagctc gtggtctcag 1020
catccctagg ccaacacccc caggattaca cagtggagaa tctcatccgc atgggtgtgg 1080
ctggcttggt cctggtggtc ctcgggattc tgctatttga ggctcagcac agccagagaa 1140
gcctacaaga tgcagccggg aggtgaacag cagagaggac aatgcatcct tcagcgtggt 1200
ggagcctcag ggacagatct gatgatccca ggaggctctg gaggacaatc taggacctac 1260
attatctgga ctgtatgctg gtcatttcta gagacagcaa tcaatatttg agtgtaagga 1320
aactgtctgg ggtgattcct agaagatcat taaactgtgg tacatttttt tgtctatgaa 1380
aaaaaaaa 1388

<210> 28
<211> 817
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 2851053CB1

<400> 28
cagcaacagc atggacatga gagtcctcgc tcagctcctg gggctcctgc tgctctgttt 60
cccaggtgcc agatgtgaca tccagatgac ccagtctcca tcctcactgt ctgcatctgt 120
tggagacaga gtcaccatca cttgtcgggc gagtcaggac attagcaatt atttagcctg 180
gtttcagcag aaaccaggga cagcccctaa gtccctgatc tatgatacat ccagtttgca 240
aagtggggtc ccatcaaagt tcagcggcag tggatctggg acagatttca ctctcaccat 300
caacagcctg cagcctgaag attttgcaac ttattactgc caacagcatc atagttatcc 360
tcttactttc ggcggaggga ccaaggtgga gatcaaacga actgtggctg caccatctgt 420
cttcatcttc ccgccatctg atgagcagtt gaaatctgga actgcctctg ttgtgtgcct 480
gctgaataac ttctatccca gagaggccaa agtacagtgg aaggtggata acgccctcca 540
atcgggtaac tcccaggaga gtgtcacaga gcaggacagc aaggacagca cctacagcct 600
cagcagcacc ctgacgctga gcaaagcaga ctacgagaaa cacaaagtct acgcctgcga 660
agtcacccat cagggcctga gctcgcccgt cacaaagagc ttcaacaggg gagagtgtta 720
gagggagaag tgtccccacc tgctcctcag ttcagcctga ccccctccca tcctttgggc 780
tctgaccctt tttccacagg gggactaccc ctattgg 817

<210> 29
<211> 936
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3238787CB1

<400> 29
tctcagtcag gacacagcat ggacatgagg gtccccgctc agctcctggg gctcctgcta 60

ctctggctcc gaggtgccag atgtgacatc cagatgaccc agtctccatc ctccctgtct 120
gcatctgtag gagacagagt caccatcact tgccgggcaa gtcagagcat tagcagctat 180
ttaaattggt atcagcagaa accagggaaa gcccctaagc tcctgatcta tgctgcatcc 240
agtttgcaaa gtggggtccc atcaaggttc agtggcagtg gatctgggac agatttcact 300
ctcaccatca gcagtctgca acctgaagat tttgcaactt actactgtca acagagttac 360
agtacccctc cgatcacctt cggccaaggg acacgactgg agattaaacg aactgtggct 420
gcaccatctg tcttcatctt cccgccatct gatgagcagt tgaaatctgg aactgcctct 480
gttgtgtgcc tgctgaataa cttctatccc agagaggcca aagtacagtg gaaggtggat 540
aacgccctcc aatcgggtaa ctcccaggag agtgtcacag agcaggacag caaggacagc 600
acctacagcc tcagcagcac cctgacgctg agcaaagcag actacgagaa acacaaagtc 660
tacgcctgcg aagtcaccca tcagggcctg agctcgcccg tcacaaagag cttcaacagg 720
ggagagtgtt agagggagaa gtgccccac ctgctcctca gttccagcct gacccctcc 780
catcctttgg cctctgaccc tttttccaca ggggacctac ccctattgcg gtcctccagc 840
tcatctttca cctcaccccc ctcctcctcc ttggctttaa ttatgctaat gttggaggag 900
aatgaataaa taaagtgaat ctttgcaaaa aaaaaa 936

<210> 30
<211> 571
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3559548CB1

<400> 30
cttgggagaa tcccctagat cacagctcct caccatggac tggacctgga gcatcctttt 60
cttggtggca gcagcaacag gtgcccactc ccaggttcac ctggtacagt ctggagctga 120
ggtgaagaag cctggggcct cagtgaaggt ctcctgcaag gcttctggtt acacctttac 180
cagtcatggt atcacctggg tgcggcaggc ccctggacaa gggcttgagt ggatggggtg 240
gatcagccct aacaatggtg acacattcta tgcacacaga ctccaggaca gagtcacctt 300
gaccacagac acatcggcga ccacaggcta catggagctg aggagcctga catctgacga 360
cacggccata tattattgcg cgagagggga ctacggtaac tctcttgacc actggggcca 420
gggaaacctg gtcactgtct cctcagcatc cccgaccagc cccaaaggtc ttcccgcttg 480
agcctctgca agcacccagc ccagatgggg aaccgtggtc aatcgcctgc cctgggtcca 540
ggggctttct tccccaaggg agccaattaa t 571

<210> 31
<211> 890
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3872741CB1

<400> 31
gtcccaacca ggacacagca tggacatgag ggtccctgct cagctcctgg ggctcctgct 60
gctctggctc tcaggtgcca gatgtgacac ccagatgacc cagtctccat cctccctgtc 120
tgcatctgtc ggagacagac tcaccatcac ttgccaggcg agtgaggacg tcatcaaata 180
tgtgaattgg tatcaacaaa aacctaggaa agcccctaaa ctcctgatcc acgatgcatc 240
caatttggaa acaggggtcc catcaaggtt cagtggaagt ggatctggga cacttttac 300
tttcaccatc agcaacctgc agcctgaaga tgttgcaaca tattactgtc agcactatgc 360

tagtcatccg ctcactttcg gcggagggac caaggtggag atcaaacgaa ctgtggctgc 420
accatctgtc ttcatcttcc cgccatctga tgagcagttg aaatctggaa ctgcctctgt 480
tgtgtgcctg ctgaataact tctatcccag agaggccaaa gtacagtgga aggtggataa 540
cgccctccaa tcgggtaact cccaggagag tgtcacagag caggacagca aggacagcac 600
ctacagcctc agcagcaccc tgacgctgag caaagcagac tacgagaaac acaaagtcta 660
cgcctgcgaa gtcacccatc agggcctgag ctcgcccgtc acaaagagct tcaacagggg 720
agagtgttag agggagaagt tgccccacct gctcctcagt tccagcctga ccccctccca 780
tcctttggcc tctgagcctt tttccacaag gggactaccc ctattgcggg tctccagctc 840
atctttcacc tcaccccgtc cttctgcttg gctttaatta tgctatgttt 890

<210> 32
<211> 928
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3981428CB1

<400> 32
ggagtcagtc tcagtcagga cacagcatgg acatgagggt ccccgctcag ctcctggggc 60
tcctgctact ctggctccga ggtgccagat gtgacatcca gatgacccag tctccatcct 120
ccctgtctgc atctgtagga gacagagtca ccatgacttg ccgggcaagt cagagcatta 180
gcacctattt aaattggtat cagcagaaac cagggaaagc ccctaagctc ttgatctatg 240
ctgcatccag tttgcaaagt ggggtcccat caaggttcag tggcagtgga tctgggacag 300
atttcactct caccatcagc agtctgcaac ctgaagattt tgcaacttac tactgtcaac 360
agagtttcaa tacccacatg tacacttttg gccaggggac acgactggag atgaaacgaa 420
ctgtggctgc accatctgtc ttcatcttcc cgccatctga tgagcagttg aaatctggaa 480
ctgcctctgt tgtgtgcctg ctgaataact tctatcccag agaggccaaa gtacagtgga 540
aggtggataa cgccctccaa tcgggtaact cccaggagag tgtcacagag caggacagca 600
aggacagcac ctacagcctc agcagcaccc tgacgctgag caaagcagac tacgagaaac 660
acaaagtcta cgcctgcgaa gtcacccatc agggcctgag ctcgcccgtc acaaagagct 720
tcaacagggg agagtgttag agggagaagt gcccccacct gctcctcagt tccagcctga 780
ccccctccca tcctttggcc tctgaccctt tttccacagg ggacctaccc ctattgcggt 840
cctccagctc atctttcacc tcacccccct cctcaagctt ggctttaatt atgctaatgt 900
tggaggagaa tgaataaata caagtgaa 928

<210> 33
<211> 762
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4635039CB1

<400> 33
ccacatctgt cctctagaga atcccctgag agctccgttc ctcaccatgg actggacctg 60
gaggatcctc ttcttggtgg cagcagtcac aggagtccac tcccaagtgc agctggtgca 120
gtctggggct gaggtgagga agcctggggc ctcagtgaag gtctcctgta aggcttctgg 180
ctacaccttc agcgaccact atattcactg ggttcgacag gcccctggac aaggacttga 240
gtggatggga tggatcaacc ctaatagtgg tggcgcaagg tatgcacagg gctttcaggg 300
cttggtcacc atgaccaggg acacgtccat cagtacagcc tacttggagc tgcgcgggct 360

gagatctgac ggctcggccg tgtacttctg tgcgagacaa accacctcgt ctcctgtagg 420
agatgctttt gatatctggg gccaagggac aatggtcacc gtctcttcag catccccgac 480
cagccccaag gtcttcccgc tgagcctctg cagcacccag ccagatggga acgtggtcat 540
cgcctgcctg gtccagggct tcttccccca ggagccactc agtgtgacct ggagcgaaac 600
ggaccagggc gtgaccgcca aaaaacttcc cacccagcca ggatgccttc gggggactgt 660
aaaccacgag cagccagctg accctgccgg gcaaaacagt gcctaacccg gaaaattcgg 720
tgaaattgcc acgtgaagca ttaaaggatt cccagccagg gt 762

<210> 34
<211> 925
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 3240710CB1

<400> 34
cttctcacca tgaggctccc tgctcagctc ctggggctgc ttatgctctg gatacctgga 60
tcgagtgcgg acattgtgtt gacccagact ccactctctc tgtccgtcac cccgggacag 120
ccggcctcca tctcctgcaa gtcttctgag agcctcctgc atactgatgg aaagacctat 180
ttgcattggt tcgtgcagaa ggcaggccag ccgccacaag tcctgatgta tgaagtctcc 240
aaccggttct ctggagtgcc agataggttc agtggcagcg ggtcaggcac agatttcaca 300
ctgaaaatca gccgggtgga ggctgaggat gttcggattt attactgcat gagaactata 360
caggttcctc caacgtggac gttcggccaa gggaccaagg tggaaatcaa acgaactgtg 420
gctgcaccat ctgtcttcat cttcccgcca tctgatgagc agttgaaatc tggaactgcc 480
tctgttgtgt gcctgctgaa taacttctat cccagagagg ccaaagtaca gtggaaggtg 540
gataacgccc tccaatcggg taactcccag gagagtgtca cagagcagga cagcaaggac 600
agcacctaca gcctcagcag caccctgacg ctgagcaaag cagactacga gaaacacaaa 660
gtctacgcct gcgaagtcac ccatcagggc ctgagctcgc ccgtcacaaa gagcttcaac 720
aggggagagt gttagaggga gaagtgcccc cacctgctcc tcagttccag cctgaccccc 780
tcccatcctt tggcctctga ccctttttcc acaggggacc tacccctatt gcggtcctcc 840
agctcatctt tcacctcacc cccctcctcc tccttggctt taattatgct aatgttggag 900
gagaatgaat aaataaagtg aatcc 925

<210> 35
<211> 1584
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4945813CB1

<400> 35
catggacctc ctgtgcaaga acatgaagca cctgtggttc ttcctcctgc tggtggcggc 60
tcccagatgg gtcctgtccc agctgcaact gcaggagtcg ggcccaggat tggtgaagcc 120
ttcggagacc ctgtccctca cctgcactgt ctctggtggc tccatcagca gttataatca 180
ctactggggc tgggtccgcc agcccccagg gaaggggctg gaatggattg ggagtatctt 240
ttatactggg aattcttact acaacccgtc cctcaagagt cgactcgcca tatcagtaga 300
cacgtccaag agccagctct ccctgaagct gagctctgtg accgccgcgg acacggccgt 360
gtattactgt gcgacagttc ctaaaacccg gtctcgacca cgtggataca cctatggtcc 420
ctttgacttc tggggccagg gaaccctggt caccgtctcc tcagcatccc cgaccagccc 480

caaggtcttc ccgctgagcc tctgcagcac ccagccagat gggaacgtgg tcatcgcctg 540
cctggtccag ggcttcttcc cccaggagcc actcagtgtg acctggagcg aaagcggaca 600
gggcgtgacc gccagaaact tcccacccag ccaggatgcc tccggggacc tgtacaccac 660
gagcagccag ctgaccctgc cggccacaca gtgcctagcc ggcaagtccg tgacatgcca 720
cgtgaagcac tacacgaatc ccagccagga tgtgactgtg ccctgcccag ttccctcaac 780
tccacctacc ccatctccct caactccacc taccccatct ccctcatgct gccacccccg 840
actgtcactg caccgaccgg ccctcgagga cctgctctta ggttcagaag cgaacctcac 900
gtgcacactg accggcctga gagatgcctc aggtgtcacc ttcacctgga cgccctcaag 960
tgggaagagc gctgttcaag gaccacctga gcgtgacctc tgtggctgct acagcgtgtc 1020
cagtgtcctg ccgggctgtg ccgagccatg gaaccatggg aagaccttca cttgcactgc 1080
tgcctacccc gagtccaaga ccccgctaac cgccaccctc tcaaaatccg gaaacacatt 1140
ccggcccgag gtccacctgc tgccgccgcc gtcggaggag ctggccctga acgagctggt 1200
gacgctgacg tgcctggcac gcggcttcag ccccaaggac gtgctggttc gctggctgca 1260
ggggtcacag gagctgcccc gcgagaagta cctgacttgg gcatcccggc aggagcccag 1320
ccagggcacc accaccttcg ctgtgaccag catactgcgc gtggcagccg aggactggaa 1380
gaaggggggac accttctcct gcatggtggg ccacgaggcc ctgccgctgg ccttcacaca 1440
gaagaccatc gaccgcttgg cgggtaaacc cacccatgtc aatgtgtctg ttgtcatggc 1500
ggaggtggac ggcacctgct actgagccgc ccgcctgtcc ccacccctga ataaactcca 1560
tgctcccccca aaaaaaaaaa aaaa 1584

<210> 36
<211> 804
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4948957CB1

<400> 36
ctgagctaca acaggcaggc aggggcagca agatggtgtt gcagacccag gtcttcattt 60
ctctgttgct ctggatctct gtcctgactg caggtgccta cggggacatc gtgatgaccc 120
agtctccaga ctccctggct gtgtctctgg gcgagagggc caccatcacc tgcaagtcca 180
gccagagtgt tttctacaac tccaacaata agaactactt agtttggtac cagcaaagac 240
caggacagcc tcctaaaatg ctcatttact gggcatctac ccgggaatcc ggggtccctg 300
accgattcag tggcagcggg tctgggacag atttcactct caccatcagc agcctgcagg 360
ctgaagatgt ggcactttat tactgtcagc aatattttac tactccgtac acttttggcc 420
aggggaccag gctggagatc aaacgaactg tggctgcacc atctgtcttc atcttcccgc 480
catctgatga gcagttgaaa tctggaactg cctctgttgt gtgcctgctg aataacttct 540
atcccagaga ggccaaagta cagtggaagg tggataacgc cctccaatcg ggtaactccc 600
aggagagtgt cacagagcag gacagcaagg acagcaccta cagcctcagc agcaccctga 660
cgctgagcaa agcagactac gagaaacaca aagtctacgc ctgcgaagtc acccatcagg 720
gcctgagctc gcccgtcaca aagagcttca acaggggaga gtgttagagg gagaagtgcc 780
cccacctgct cctcagttcc aggg 804

<210> 37
<211> 878
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 4949649CB1

<400> 37
gcctgagagc acttagcgtt catgagtgtc cccaccatgg cctggatgat gcttctcctc 60
ggactccttg cttatggatc aggagtggat tctcagactg tggtgaccca agagccatcg 120
ttatcagtgt cccctggagg gacagtcaca ctcacttgtg gcttggcctc tgactcagtc 180
tctactaatt tcttccccac ctggtaccag cagaccccag gccaggctcc acgcacgctc 240
atctacagca caagcactcg ctcttctggg gtccctgatc gtttctctgg ctccatcctt 300
gggaacaaag ctgccctcac cattacgggg gcccaggcag atgatgaatc tgattattac 360
tgtgcgcttt atatgggcag tggcatttcg gtattcggcg gagggaccaa ggtgaccgtc 420
ctaggtcagc ccaaggctgc cccctcggtc actctgttcc cgccctcctc tgaggagctt 480
caagccaaca aggccacact ggtgtgtctc ataagtgact tctacccggg agccgtgaca 540
gtggcctgga aggcagatag cagccccgtc aaggcgggag tggagaccac cacaccctcc 600
aaacaaagca acaacaagta cgcggccagc agctatctga gcctgacgcc tgagcagtgg 660
aagtcccaca gaagctacag ctgccaggtc acgcatgaag ggagcaccgt ggagaagaca 720
gtggccccta cagaatgttc ataggttctc aaccctcacc ccccaccacg ggagactaga 780
gctgcaggat cccaggggag gggtctctcc tcccacccca aggcatcaag ccctttctccc 840
tgcactcaat aaaccctcaa taaatattct cattgtca 878

<210> 38
<211> 1680
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID NO: 5500302CB1

<400> 38
ctgggcgctg cggcggcagg aggaggacgg ggaaggacgg agccgagccg cggctgcctc 60
cctcgctcac tccctcgcgc actcgcccgc cccctccctc cctcccctcc cttccccggc 120
cccggctctg gccccggccc attcgctgtt gggtcttctg ctagggagga tgtcgggttc 180
gtcgctgccc agcgccctgg ccctctcgct gttgctggtc tctggctccc tcctcccagg 240
gccaggcgcc gctcagaacg ctgggtttgt caagtcgccc atgtcagaaa ctaagctcac 300
gggggacgcc tttgagctgt actgtgacgt ggtcgggagc cccacgccag agatccagtg 360
gtggtacgca gaagtcaacc gggcagagtc tttcagacag ctgtgggacg gtgctcggaa 420
gcgccgtgtc accgtaaaca ccgcctacgg gtcaaacggc gtgagtgtgc tgagaataac 480
ccggctcacc ttggaggact ctgggactta cgagtgcagg gccagcaacg accccaagag 540
gaatgacttg aggcaaaacc cctccataac atggattcga gcccaggcca ccataagcgt 600
ccttcagaag ccaaggattg tcaccagtga agaggtcatt attcgagaca gccctgttct 660
ccctgtcacc ctgcagtgta acctcacctc cagctctcac acccttacat acagctactg 720
gacaaagaat ggggtggaac tgagtgccac tcgtaagaat gccagcaaca tggagtacag 780
gatcaataag ccgagagctg aggattcagg cgaataccac tgcgtatatc actttgtcag 840
cgctcctaaa gcaaacgcca ccattgaagt gaaagccgct cctgacatca ctggccataa 900
acggagtgag aacaagaatg aagggcagga tgccactatg tattgcaagt cagttggcta 960
cccccaccca gactggatat ggcgcaagaa ggagaacggg atgcccatgg acattgtcaa 1020
tacctctggc cgcttcttca tcatcaacaa ggaaaattac actgagttga acattgtgaa 1080
cctgcagatc acggaagacc ctggcgagta tgaatgtaat gccaccaacg ccattggctc 1140
cgcctctgtt gtcactgtcc tcagggtgcg gagccacctg gccccactct ggcctttctt 1200
gggaattctg gctgaaatta tcatccttgt ggtgatcatt gttgtgtatg agaagaggaa 1260
gaggccagat gaggttcctg acgatgatga accagctgga ccaatgaaaa ccaactctac 1320
caacaatcac aaagataaaa acttgcgcca gagaaacaca aattaagtac tgcttacaat 1380
atctttaggt tcctgaaact ggtggcaaca tgacctgcta aaattttctg cttggacctc 1440
tttggttctc tccccttca agtgagcaac accacaatga ctgtctaaag catgccttat 1500
ttagcctctc ctgtaagggt gatctagcca ggtacatttt aaacaatgct tcagtgtaga 1560
aggtgtaaac tattttgggc ttgatgtgct gtgaatgttg ctttttttt ccttggtaaa 1620

atatttaaat tagaagtgaa aaggtcctct gaggatcaga ttcatgcatg cgccattttt 1680